Croatian Journal of Philosophy

Aesthetics and Philosophy of Art

IRIS VIDMAR JOVANOVIĆ

ELISA CALDAROLA

ALEXANDRA DIAS FORTES

POLONA TRATNIK

ELENA ABATE

DAVID COLLINS

ANDREW CORSA

PHILIP MILLS

RAFE MCGREGOR

BORAN BERČIĆ

Article

DAVID PEREPLYOTCHIK

Book Reviews

Vol. XXI · No. 61 · 2021

Croatian Journal of Philosophy

1333-1108 (Print)
1847-6139 (Online)

Co-published by
"Kruzak d.o.o."
Naserov trg 6, 10020 Zagreb, Croatia
fax: + 385 1 65 90 416, e-mail: kruzak@kruzak.hr
www.kruzak.hr

and

Institute of Philosophy
Ulica grada Vukovara 54/IV, 10000 Zagreb, Croatia
fax: + 385 1 61 50 338, e-mail: filozof@ifzg.hr
www.ifzg.hr

Available online at http://www.ceeol.com and www.pdcnet.org

CROATIAN
JOURNAL
OF PHILOSOPHY

Vol. XXI · No. 61 · 2021

Aesthetics and Philosophy of Art

Article

Book Reviews

Introduction

The latest issue of the Croatian Journal of Philosophy *presents nine papers dedicated to topics in philosophy of art and aesthetics. This is the second time that an issue of Croatian has been dedicated to these topics and we hope it will invite further submissions from aestheticians and philosophers working on art. We originally issued a CFP in the spring of 2020 and over the course of the following months we had a great flow of highly interesting papers coming in. Selecting the papers that are included in this issue was not an easy process, and we would like to thank all the contributors and all the reviewers for their efforts and contributions.*

Art, and art-related practices are changing and developing in unpredictable ways, with artists constantly finding new ways of putting art to the service of engaging with our everyday experience. The selection of papers here reflects some of these engagements and testifies to how philosophically intriguing they are. We are very fortunate for managing to come up with a great diversity of papers. They range from those that address some traditional concerns of philosophical aesthetics, such as expressivism in music or art's capacity to provide knowledge, to more contemporary ones, including a paper on fashion and two papers on architecture. Diversity is also reflected in varieties of approaches to aesthetics, such as the one combining war aesthetics and literary theory or the one dedicated to exploring the notion of art in the context of political criticism of consumerism. Some of our contributors show how analytic and continental philosophy unite, for example, in unraveling the value of poetry or bringing together art forms and life via the figure of Wittgenstein. We hope our readers will enjoy this diversity and feel inspired to respond with their papers.

Here is a brief overview of the papers.

Elisa Calderola offers an impressive and much needed analysis of the relation between architectural objects and their sites, and she does so by employing the framework of the site-specific art. She fortifies this analysis with a debate on the categorization of artworks as site-specific. The paper is ridden with examples of architecture from around the world, offering insightful information on the context of their creation and exposing philosophical implications of that context. As Calderola argues, her theory provides a unified account of site-specific art and architecture and it illuminates the growing reference of architecture being closely site-focused.

Alexandra Dias Fortes' paper continues the topic of architecture. In her paper, architecture is positioned as a background from which the author explores urban life modes. Fortes offers an insightful analysis of Aldo Rossi's architecture, Georg Trakl's poetry and Wittgenstein's philosophy, thus exploring various ways in which art and philosophy intersect in our attempts to understand the material circumstances of our lives.

Polona Tratnik explores manners in which art can be put to the service of politically engaged criticism of consumerism. She does so by engaging with Fredric Jameson's theory, who, as Tratnik argues, fails to address the problem because of his sole concern with representational art. To properly understand critical or political art, Tratnik argues, we should focus on performative art. Such art is set outside of space traditionally designed for art and occupies the space intended for consumerism, which enables it to be critical.

The merging of art and the everyday is further pursued in Elena Abate's paper. Abate explores fashion as an aesthetic form of art, developing her account against Wittgenstein's aesthetic conception. Abate sees fashion as a point of contact between the grammar of language and socially encoded aesthetic responses which offer an individual the possibility to constantly reinvent himself creatively.

David Collins engages with a debate on the musical expression of emotions, focusing mainly on Stephen Davies and Jerrold Levinson's theories. While the two are considered to be rival theories, Collins offers an account of "expressing," which neutralizes the alleged opposition between the two theories. A central aspect of the paper is its treatment of Collingwood's theory of artistic expression, which Collins relies upon to explain the relation between music and emotions.

Andrew Corsa engages in a debate regarding literature's capacity to instill knowledge and moral lessons in the readers. Corsa approaches this question by focusing on how a novel might help individuals create a more meaningful life by redirecting their future-directed personal narratives.

The literature's cognitive value is further taken up in a paper by Philip Mills, who defends poetry against Austin-inspired criticism which sees it is "forceless". Mills sets out to explore poetry's ways of gaining linguistic, philosophical and political force and argues that the force of poetry resides in its capacity to change our ways of seeing. Rephrasing Austin, Mills concludes that poetry has the power to do things to words and by doing so, to transform the world.

Another paper dedicated to explaining the cognitive value of literature is Rafe McGregor's. McGregor develops a literary aesthetics of war crime and does so by examining the phenomenon of moral immunity in military memoir. His focus is on unjust wars, and he identifies three literary devices in which moral immunity is achieved: literary irresponsibility, ethical peerage and moral economy. In the second part of the pa-

per McGregor relates his findings to the value interaction debate on the one hand, and to the ethics of reading on the other. The literary analysis of novels provided in support of the argument is a treat in its own right.

Boran Berčić approaches art from the metaphysical point of view, exploring manners in which art can represent the impossible. The central debate that Berčić considers is one between possibilism and impossibilism, and he ultimately develops an account in favor of possibilism. The most engaging aspect of the paper is centered around five possible ways in which these issues can be asked and Berčić provides an interesting set of examples pertaining to various forms of art to unravel ways in which art might represent what is impossible.

IRIS VIDMAR JOVANOVIĆ

Croatian Journal of Philosophy
Vol. XXI, No. 61, 2021
https://doi.org/10.52685/cjp.21.1.1
Received: August 20, 2020
Accepted: December 7, 2020

Architecture and Sites: A Lesson from the Categorisation of Artworks

ELISA CALDAROLA
University of Padua, Padua, Italy

Several contemporary architects have designed architectural objects that are closely linked to their particular sites. An in-depth study of the relevant relationship holding between those objects and their sites is, however, missing. This paper addresses the issue, arguing that those architectural objects are akin to works of site-specific art. In section (1), I introduce the topic of the paper. In section (2), I critically analyse the debate on the categorisation of artworks as site-specific. In section (3), I apply to architecture the lesson learned from the analysis of the art debate.

Keywords: Philosophy of art; site-specific architecture: site-specific art; Snøhetta; Rem Koolhaas.

Several contemporary architects have designed architectural objects that are closely linked to their particular sites. An in-depth study of the relevant relationship holding between those objects and their sites is, however, missing. This paper addresses the issue, arguing that those architectural objects are akin to works of site-specific art. In section (1), I introduce the topic of the paper. In section (2), I critically analyse the debate on the categorisation of artworks as site-specific. In section (3), I apply to architecture the lesson learned from the analysis of the art debate.

1. *A look at contemporary architecture*

In current architectural practice, there is growing interest in designing architectural objects—a term that I use to refer to all sorts of built structures (see Fisher 2015)—that are closely site-linked, as testified by the work of, e.g., Peter Zumthor, Róisín Heneghan (co-founder of Heneghan Peng Architects along with Shi-Fu Peng), and Kjetil Thors-

en and Craig Dykers (founders of the architectural practice Snøhetta).[1] Moreover, the design of public spaces and landscapes is gaining importance, as recent projects by Bernard Lassus, Kathryn Gustafson, and Neil Porter show. Kjetil Thorsen declared in a 2015 interview: "…in the next 20 years, we'll see more of a shift towards public spaces which are between […] built objects and lesser focus on the *iconic building*" (Forbes 2015: 126, my italics; see also Spens 2007).[2]

Leslie Sklair (2006: 25–33) explains that a building is usually labelled 'iconic' when it is famous as well as symbolically and aesthetically relevant—St. Peter's Basilica in Rome, for instance, is world-famous, and charged with symbolic meaning (it is the most prominent building in the Vatican City and that most often used by the pope, thus being a symbol of Roman Catholicism itself), and a masterpiece of Renaissance art. As Sklair argues (33–43), a number of contemporary buildings described as iconic are aesthetically remarkable, have been created by famous architects and/or made famous through images in the media, and have been financed by global corporations, rather than by state and/or religious bodies, with the goal of making cities "easily recognizable for purposes of commerce and civic pride. […] Those driving urban boosterism deliberately attempt to create urban architectural icons in order to draw tourists, convention and mega-event attendees with money to spend and the images they project are directed to this end" (38). Buildings of this kind, Sklair suggests, function as symbols of globalized capitalism itself (33). An example is Frank Gehry's Disney Concert Hall (2003), which contributed to changing the image of downtown Los Angeles, inserting into an unremarkable area usually shared between office workers and homeless people a monumental, asymmetrical building, with curved, stainless steel façades, for concertgoers and tourists (34). Interestingly, a consequence of the contemporary trend in iconic architecture-making is that in cities as diverse as, for example, Los Angeles, Bilbao, and Dubai we can find buildings that serve the same chief inter-

[1] The kinds of relationships between architectural objects and their sites that I shall take into consideration in what follows by no means exhaust the realm of the relationships that can obtain between an architectural object and its site. As it shall emerge throughout the paper, my interest lies exclusively in a feature of architectural objects—the fact that their spatial extension encompasses their sites, in a way—and, more specifically, in architectural objects that articulate some content about their context of production through the sites they encompass. It is this latter, peculiar kind of relationship that I have in mind when I write of architectural objects that are *closely linked* to their particular sites. I shall not concern myself with, e.g., architectural objects built with construction techniques and/or with materials that are specific to their sites.

[2] In what follows, my analysis shall focus only on buildings, rather than, more generally, on all sorts of architectural objects, merely for the sake of simplicity. I assume that the claims I shall make with reference to buildings can be applied to all sorts of architectural objects. Whether my claims also apply to non-built objects, such as public spaces and landscapes, is a separate question, which I shall not address here.

est (i.e., to change a city's image in order to boost its economy) and that, in order to attract consumers, use the same strategy, which consists in creating mega-structures with unprecedented shapes that stand out against their background—in addition to Gehry's Disney Concert Hall, for example, consider also his Guggenheim Museum in Bilbao (1997) and Dubai's Burj Khalifa (2009) by Skidmore, Owings, and Merrill. A number of contemporary iconic buildings, then, seems to be have been produced mainly with the goal of changing cities' images in order to "commodify the urban experience" (Sklair 2012: 352), to the detriment of their fitting into their natural, architectural, social and cultural contexts—which is instead the focus of the practice of building architectural objects that are closely linked to their particular sites.[3]

An example of the latter approach is the Norwegian National Opera and Ballet (2008) in Oslo, by Snøhetta. The building, which hosts three state-of-the-art theatres of different sizes and provides work-space for around 600 employees, has monumental proportions, but merges harmonically with its surroundings thanks to "a 'carpet' of horizontal and sloping surfaces" (Snøhetta 2008: n. p.) laid out on top of it, whose form relates to the cityscape. In particular, "Viewed from the Akershus castle and from the grid city the building creates a relationship between the fjord and the Ekerberg hill to the east. Seen from the central station and Chr. Fredriks square the opera catches the attention with a falling which frames the eastern edge of the view of the fjord and its islands. The building connects city and fjord, urbanity and landscape" (Snøhetta 2008: n. p.). The Opera House establishes connections not only at landscape level, but also at social level: it provides a public plaza and a foyer that are freely accessible to the public, as well as shops and restaurants. Numerous people spend time in those areas, which have effectively provided accessible, well-designed public spaces to Osloites, in a previously neglected part of the city.

The main goal of this paper is to look deeper into architectural objects that are closely linked to their particular sites. To my knowledge, literature on this topic is rather limited and polarized: on the architectural theory side, there are volumes which are less preoccupied with providing a general understanding of those objects than with investigating the views of particular architects engaged in the practice of producing them (e.g., Spens 2007; Forbes 2015; Aldallal et al. 2016), while on the philosophy side Fabio Bacchini (2017) has provided an insightful, although ultimately inadequate account, as I shall argue in section (3). Debate on artworks that are closely linked to their particular sites is, instead, more developed (see, e.g., Crimp 1986; Crow 1996; Coles 2000; Kaye 2000; Suderburg 2000; Kwon 2002; Gaiger 2009; Rugg 2010): this suggests the strategy of looking into art-theoretical discussions first, and then at-

[3] Closely site-linked architecture and landscaping can of course be tools for urban boosterism too, as shown by projects like West 8 and DTAH's Toronto Waterfront Revitalisation Initiative (2006) and The High Line, Manhattan (2009) by James Corner Field Operations, Diller Scofidio + Renfro and Piet Oudolf.

tempting to make analogies between artworks, in general, on the one hand, and architectural objects, on the other—this strategy has already been pursued by Bacchini (2017) but, as I shall show in section (3), his reading of the debate on closely site-linked art lacks critical insight. An intuitive reason supporting this strategy is that various architectural objects are usually considered artworks: an analysis of architectural objects that are closely linked to their particular sites could thus simply consist in the application to the particular case of artistic architecture of a more general view concerning all artworks. One, however, might object that it would be preferable to have an account not just of *artistic* architectural objects that are closely linked to their particular sites, but of *all* architectural objects that are closely linked to their particular sites. As I shall explain in section (3), however, this problem does not arise, because my claims concerning the relationship between art- and architectural objects and their particular sites can be applied also to non-artistic artifacts, in general, and non-artistic architectural objects, in particular. In the rest of the paper, I shall proceed as follows: in section (2), I shall look into the debate on how artworks relate to their particular sites, and I shall dispel a confusion emerging from it, putting forward an argument for distinguishing between the macro-category of *sited* artworks and the sub-category of *site-specific* artworks. In section (3), I shall apply to architecture the lessons drawn from the scrutiny of artworks, thereby providing an original account of how architectural objects closely relate to their particular sites and underlining the points of contact and the divergences between Bacchini's (2017) account and mine.

2. *Sited art and site-specific art*[4]

All particular artworks and all instantiations of multiply instantiable artworks are *physically located* somewhere. For instance, the *Mona Lisa* is displayed in the Louvre and any execution of the *Ninth Symphony* happens at a specific time and place. The physical location of any artwork or artwork-instantiation is relevant to one's experience of it in so far as it allows for making the artwork, or artwork-instantiation, perceivable in a way that respects its author's "sanctions" i.e. "publicly accessible actions and communications" (Irvin 2005: 315)—such as presenting a work within a certain context, giving the work a certain title, offering an artist's statement about the work, instructing curators on how, for example, to display the work (Irvin 2005: 319–320)—concerning how the boundaries of a given artwork are to be fixed, what features of the work are relevant to interpreting it, what genre the work belongs to, and whether the work has a particular feature as an artwork or not (Irvin 2005: 315–316).[5] For instance, for a proper experience of the *Mona Lisa*,

[4] This section partially relies on arguments put forward in Caldarola 2020: ch. 3.

[5] As Irvin explains (2005: 221–222), sanctions are not necessarily explicitly spelled out, or even established, by artists: they can also result from an artist's

the Louvre room where the painting hangs must be sufficiently illuminated, because, by producing a painting within the Western tradition of picture-making, Leonardo sanctioned that in order to experience the *Mona Lisa* we need to be able to perceive what the painting itself depicts.

On the other hand, only some artworks and artwork-instantiations are such that their *media* are constituted by certain physical objects or events *plus* portions of the physical environments where they are located or take place—as can be inferred from their makers' sanctions.[6] I shall call those artworks and artwork-instantiations 'sited'. A sited artwork-instantiation is, for instance, the instantiation of Auguste Rodin's *Les Bourgeois de Calais* (1889) which is installed in a square in Calais—there are eleven more instantiations of the statue, some shown in museums and others installed in public spaces. The Calais instantiation of the statue celebrates six fourteenth century citizens of Calais and, thanks to its public location in Calais, it manifests the city's pride for its past. Understanding that it manifests the city's pride for its past is integral to one's complete appreciation of it, which is why it is reasonable to claim that the Calais square where the statue is installed is part of the medium of this artwork-instantiation.[7]

The term 'site-specific' first appeared in the art-jargon in the late 1970s/early 1980s[8] and art-theoretical literature usually considers site-specific art as a contemporary phenomenon that emerged in the 1960s (see, e.g., Crimp 1986; Foster 1996; Meyer 2000; Coles 2000; Kaye 2000; Kwon 2002). According to the literature, the goal of most site-specific artworks and artwork-instantiations is either to make the spectator aware of her presence in the physical space and of aspects of her experience of it, or to criticize certain institutions by re-configuring their physical sites through actions and installations (see especially Kwon 2002 and, for an amendment to her view, Gaiger 2009). I shall clarify this with two examples. A 1970s work that invites spectators to focus on their experience of the physical space is, for instance, Nancy Holt's *Sun Tunnels* (1976). This sculptural installation located in the Great Basin Desert in Utah consists of four concrete cylinders arranged in an open cross format and aligned to frame the sun on the horizon during the summer and winter solstices. The work can be described as a device for experiencing the sun's light at certain times of the year in a particular way and, given how the cylinders are positioned, it can only

subscription to conventions that are established in the community where there is a tradition which grounds the artist's practice.

[6] A medium is any resource manipulated in some way in order to convey content and/or make salient some properties of an object (see Lopes 2014: ch. 7).

[7] While the property of being physically located somewhere is an *extrinsic* property of all particular artworks and artwork-instantiations, the property of being sited is an *intrinsic* property of only some particular artworks and artwork-instantiations (see Marshall and Weatherson 2018).

[8] See Robert Irwin's *Being and Circumstance* (1985) and the debate on Richard Serra's *Tilted Arc* (1981–1989).

function in the very site where it is installed. On the other hand, a work that criticizes an institution by occupying its site is Mierle Laderman Ukele's *Hartford Wash* (1973): in her performance at the Wadsworth Atheneum—a museum in Hartford, CT—the artist engaged in ordinary house-maintenance tasks, such as washing the floors, the goal being to raise awareness of the under-representation and marginalization of female artists in museums like the Wadsworth Atheneum (see Steinhauer 2017).

As the above examples show, site-specific artworks and artwork-instantiations are, certainly, sited artworks and artwork-instantiations: their *media* consist of particular material objects or events *plus* portions of the environments in which they are located or take place. But are they *just* sited artworks and artwork-instantiations? This open question has generated some confusion. On the one hand, literature on site-specificity has focussed on the *reasons* prompting contemporary artists to produce sited artworks and artwork-instantiations—mainly, their interest in having the public actively explore portions of the physical space and in criticizing institutions by re-configuring their physical sites, as I have mentioned above. This might support the hypothesis that what distinguishes contemporary site-specific art from the broader category of sited art is simply the particular motives of contemporary artists engaged in the production of sited works and works-instantiations, rather than aspects of the works and works-instantiations themselves (to my knowledge, however, this view has, as yet, not been explicitly defended). On the other hand, some art theorists and historians have focussed on 'siting' when considering, for example, some Renaissance and Baroque artworks, establishing links with literature on contemporary site-specific art and exploring the hypothesis that the realm of site-specific art might extend well beyond contemporary art production (see, e.g., Gillgren 2011 and 2017, as well as some essays contained in Gillgren and Snickare 2012). My view is that a more careful analysis of works and work-instantiations usually described as cases of site-specific art allows us to understand that it is appropriate to distinguish the sub-category of site-specific art from the macro-category of sited art, but for reasons other than those that might be inferred from the literature on contemporary site-specific art: site-specific artworks and artwork-instantiations should be distinguished from sited artworks and artwork-instantiations not merely in virtue of the motives that guided their production, but also in virtue of the fact that their sites are *essential* for them to convey content that concerns the historical and/or social and/or cultural context of their production. In the following paragraphs, I shall support this view with the analysis of three widely recognized examples of site-specific art and three examples of sited art that isn't usually considered site-specific.

Let us first consider paradigmatic cases of site-specific art. I believe that *Sun Tunnels* is not just a work that uses its site to produce an in-

tense experience of physical space in the spectator, but also a work that uses its site, crucially, to articulate content about the cultural context of its production: through its collocation in the desert and significant size, it antagonizes the art-gallery system of the 1970s, which promoted the commercialization of artworks, understood as easily movable commodities. The work distances itself from the prevailing attitude in its cultural context of production, thereby implying a critical commentary about it (see Williams 2011).

Let's now go back to *Hartford Wash*. It is an instantiation of a work of performance, whose art medium incorporates the Wadsworth Atheneum's site: if the artist had been asked to wash the floors of Hartford's railway station, instead of those of the museum, our experience of her actions wouldn't have qualified as the experience of her performance work. The performance was, in part, about the historical and cultural context of its production: namely, it criticized views about women artists held in the art-world in the 1970s. Furthermore, the fact that the performance was situated in a museum was crucial for it to convey its context-related content: in order to criticize the museum's marginalization of women artists, the performance had to show a woman artist washing the floors *of a museum*—it couldn't show a woman artist washing, say, the floors of a railway station.

Finally, let us analyse another instantiation of a site-specific artwork: Katarina Fritsche's sculpture *Hahn/Cock*, installed in London's Trafalgar Square between 2013 and 2015. The statue, representing a blue cockerel (an ironic symbol of male power), was to be experienced as a critical commentary towards the numerous symbols of male, imperialistic power constituted by the other statues installed in the square (first of all, that of Admiral Nelson). Furthermore, the statue expressed contemporary discomfort towards the nineteenth century British Empire, qualifying as a work instantiation that was also about the historical and cultural context of its production (i.e., a work instantiation that conveyed a contemporary view on former European imperialism) (see Barnett 2013). Lastly, if the work instantiation had not been situated in Trafalgar Square it could not have conveyed its context-related content: if the cockerel had been installed in a square that hadn't presented statues that are powerful symbols of imperial power, it couldn't have expressed contemporary discomfort towards the British Empire of the nineteenth century.

So far, I have shown that the sites of site-specific artworks and artwork-instantiations are *essential* for them to convey content that is about the historical and/or social and/or cultural context of their production. This observation allows us to draw a first distinction between site-specific art and some other cases of sited art. Consider, for instance, the Calais instantiation of *Les Bourgeois de Calais*: this work-instantiation testifies to the artistic style of Rodin, a late nineteenth century French sculptor, as well as to its age's technical achievements in bronze cast-

ing, although, to my knowledge, it is not about late nineteenth century French history, culture or society. More generally, while all artworks and artwork-instantiations, to some extent, *testify* to their respective historical and/or social and/or cultural contexts, not all artworks and artwork-instantiations are such that they have been sanctioned to convey a view about those contexts. Consequently, site-specific artworks and artwork-instantiations, qua objects that have been sanctioned to convey a view about their contexts of production, can be distinguished from those sited artworks and artwork-instantiations which, like the Calais instantiation of *Les Bourgeois de Calais*, have *not* been sanctioned to convey a view about their contexts of production.

The next step in my argument consists in showing that some sited artworks or artwork instantiations that *are* about the historical and/or social and/or cultural context of their production do not function like paradigmatic cases of site-specific artworks and artwork-instantiations such as *Sun Tunnels*, *Hartford Wash* and *Hahn/Cock*. To support this view, I shall give two examples. One, the *Ara Pacis Augustae* (9 BC), is an altar which was sited with respect to its original location, the Campo Marzio in Rome—an area dedicated to the celebration of the conquests of the Roman Empire (the work is now installed in a different area of Rome, by the bank of the Tiber). The altar—which, by depicting the Emperor, his family and other prominent figures officiating a procession to celebrate the *pax augustea*, pays tribute to the exceptional season of peace Emperor Augustus brought to the Roman Empire, and is therefore about the historical context of its production—had been sanctioned to be experienced among other monuments celebrating war victories in the Campo Marzio, which created an appropriate framework for stressing the uniqueness and relevance of Augustus' achievement. This might suggest that there is no difference between the *Ara Pacis* and, for example, Fritsche's *Hahn/Cock*: both are sited and about the historical/social context of their production. However, there is an important difference between the two works: unlike the latter, the *Ara Pacis* conveys its context-related content even if it has been removed from its original location (as mentioned, the work, originally in the Campo Marzio, now stands by the bank of the Tiber). In particular, the work could convey its context-related content even if it were removed from the city of Rome and transported to, say, Beijing. This is because the part of its content that concerns the celebration of the Augustan Age is conveyed by the scenes depicted by the sculptural reliefs, and not by the original site of installation of the artwork (i.e., the Campo Marzio in Rome). The original location merely emphasized the unicity of August's achievement, contrasting the *Ara*—a monument celebrating peace—with other monuments celebrating conquests.

My second example of sited art that is about its context of production, although it differs from typical works of site-specific art, is Paolo Veronese's *Nozze di Cana* painting (1563), originally installed in the dining room of the San Giorgio Basil in Venice and now displayed in the

Louvre. This work used to be situated in the basil's dining room, since it represents a scene that can be imagined as extending into the dining room's environment. Furthermore, the work is also, in part, about the social and cultural context of its production, since it represents a sixteenth century environment and several celebrities of the time (e.g., Titian and Veronese himself). Like the *Ara Pacis*, however, *Nozze di Cana* differs from *Sun Tunnels*, *Hartford Wash* and *Hahn/Cock* in that the site where it was once situated was not essential in conveying the part of the work's content that concerns the social and cultural context of its production. Indeed, the work is about sixteenth century Venetian society and culture, because it depicts a particular environment and particular individuals, and not because it incorporated the dining room within its artistic medium.

To sum up and conclude this section, I argue that site-specific artworks and artwork-instantiations differ from other sited artworks and artwork-instantiations in that they use their sites, crucially, to convey content about their contexts of production. To my knowledge, the realm of site-specific art encompasses only some *contemporary* sited artworks and artwork-instantiations; however, it might emerge that some works and work-instantiations from previous epochs also qualify as site-specific art. Art-theoretical literature on site-specific art *might* therefore be right in focusing exclusively on contemporary sited artworks, but, as I have shown, it is inadequate in that it has not thoroughly investigated the reasons underpinning the description of a work or work-instantiation as site-specific. My analysis unveils the peculiarities of site-specific art. In the next section, guided by my view on sited and site-specific art, I shall attempt to provide a deeper understanding of architectural objects that enjoy close relationships with their sites.

3. *Sited architecture and site-specific architecture*

Let us finally turn to architectural objects that are closely linked to their particular sites. In the previous section, I introduced the notion of *sited* artwork or artwork-instantiation, which is such that its medium is constituted by certain physical objects or events *plus* portions of the physical environment where it is located or takes place. All architectural objects, I submit, are *sited*. On this point, I agree with Bacchini, who argues that "buildings are constitutively located in a certain place" (2017: 86).[9] Before summarizing and endorsing Bacchini's arguments in support of this claim, I would like to set the ground for addressing a preliminary objection that might be raised against my approach towards architectural objects.

One might object that, since I shall argue for an analogy between artworks and artwork-instantiations, on the one hand, and architectural objects, on the other, my claims will apply exclusively to those architectural objects that qualify as art (if there are any). However,

[9] See footnote 1 above.

this is false. As we have seen in section (2), my view on sited and site-specific art revolves around a particular conception of their *medium*: sited artworks and artwork-instantiations are such that their medium is constituted by certain physical objects or events *plus* portions of the physical environment where they are located or take place (i.e., their sites), while site-specific artworks and artwork-instantiations are sited artworks and artwork-instantiations which use their sites, crucially, to convey content about their contexts of production. It is important to note that, as Dominic Lopes (2014: ch. 7) argues, it is not just art objects that have media: all resources manipulated in some way in order to convey content and/or make salient some properties of an object qualify as media. Street signs, supermarket leaflets, and IKEA mugs all have media. I believe that all architectural objects, too, have media, because they are constituted by materials that make salient some of their properties (I shall develop on this below). What distinguishes an art-medium from a non-artistic medium, Lopes argues, is that only the former is the focus of an art-appreciative practice, but this issue shouldn't concern us here.

Let us now look at Bacchini's arguments. He begins with the observation that "we may doubt whether a specific building would still be the same if we moved it to another location" (2017: 88). For instance, we tend to think that changing Notre Dame's location from Paris to Las Vegas would imply altering one of the essential properties of the church (i.e., that of being located in Paris). On looking for an explanation for this kind of intuition, Bacchini discards various prima facie hypotheses: what grounds the intuition is not the fact that buildings are originally conceived by their makers as permanently located, because we can imagine that we would still hold the intuition about Notre Dame even if we knew for sure that its makers didn't conceive of it as permanently located (88); the fact that buildings have foundations doesn't ground the intuition either, because we hold it also for buildings that, like the Parthenon, don't have foundations, and because we don't hold similar views about trees, whose roots, however, play the same role that foundations play for buildings (91); finally, the intuition is not grounded in the mere fact that buildings are difficult to move because (a) that is an accidental property of buildings, while we think that their being located in a certain place is one of their essential properties, and (b) there are other objects that are difficult to move (e.g., rather big meteorites) about which, however, we don't hold the same intuition (90).

Bacchini's next step is to clarify that the essential property that we think is altered by changing the position of a building is not the property of being located at a particular latitude and longitude, but of being surrounded by a specific physical context. He supports this claim by observing, on the one hand, that we wouldn't think that, for instance, Notre Dame had changed one of its essential properties if it had been moved, along with the whole city of Paris, to Clark County, Nevada. In

particular, it is not relevant for us to know whether the relocated city is identical to Paris, while it is relevant to know that the outer context of the church is the same as in Paris (91–92). On the other hand, Bacchini observes that if the church of Notre Dame were left in its original location, but "the whole town of Paris around the Cathedral" (92) were replaced by the town of Las Vegas, we would hold that the church would have lost its essential property of being situated in a certain location.

Having made this further point, Bacchini claims that what explains the intuition that the property of being located in a certain physical context is an essential property of buildings is the fact that, for any given work of architecture, "usually there is no basis for ruling out any extrinsic contextual feature of the work as inessential" (95), since "it is manifest that we usually allow among the constitutive properties of edifices much more intrinsic properties than just those indicated in plans (for example, materials, interaction with sunlight, shadows)" (94).

I endorse Bacchini's arguments and shall try to clarify them with some examples. To begin with, let us look at Frank Lloyd Wright's *Fallingwater* (1938)—a building that was famously conceived with the goal of fitting it to the natural environment where it is situated. Among the constitutive properties of this building there are, for instance, the property of looking as if it literally incorporated a waterfall (while it has actually been built *upon* one) and the property of being embedded among trees with particular shapes and colours. This is evident both from what we know about Lloyd Wright's project for the building (see, e.g., Gibson 2017; PPG 2020) and from our experience of it: it is integral to our appreciation of the building's structure that we focus on how masterfully it has been placed above the waterfall and that we notice, for instance, that the colour of its exterior is similar to the colour of the dying rhododendrons that can be seen around the house in the fall.

As a second example, let us consider a contemporary iconic work of architecture such as Rem Koolhaas' CCTV tower (2013) in Beijing. Even though this is an extraordinary-looking building—a folded tower of huge proportions—which draws attention to itself more than to its surroundings, it would be mistaken to claim that we shouldn't take into consideration some of its relationships with its location when experiencing it. For instance, the architecture of the building—which has been described as "not phallic, but vaginal" (de Muynck 2004: n. p.)—is capable of attracting people around itself and is better appreciated when seen against the background of the many nondescript or whimsical high-rises and skyscrapers punctuating the landscape of contemporary Beijing, which are like "needles" that "collect their own little pathetic communities while breaking down the larger community around them" creating "isolation right in the center of the city" (Koolhaas 2004: 465).

Finally, let us focus on a kind of building that has had numerous instantiations in former East Germany, as well as in former West Ger-

many and the Netherlands: the *Plattenbau*. *Plattenbauten* made of large, prefabricated concrete slabs were mainly employed for cheap, public housing projects, often occupying large portions of settlements on the outskirts of large cities. Consequently, some of those settlements look quite similar to each other and one might think that relocating a certain *Plattenbau* from, say, Lichtenberg (a Berlin neighbourhood) to, say, Marzahn (another Berlin neighbourhood) wouldn't alter any essential property of the building. Following Bacchini, however, it can be objected that, even in this case, in which a building is moved to a neighbourhood that is very similar to its original one, we cannot know for certain whether, in appreciating the building, it would be right to disregard, say, the particular way in which the light hits it at dusk in its Lichtenberg location, which is in no way equal to the light that hits it at dusk in its Marzahn location.

Bacchini concludes by addressing a problem raised by his view: it is not clear why, while the original physical context of a building tends to change, sometimes to a great extent, over time, we don't seem to hold the view that the building loses any of its essential properties as a consequence of those changes in its physical context. The portion of Paris around Notre Dame, for instance, has changed significantly since the thirteenth century, but this doesn't seem to concern us (95–98). To make sense of this phenomenon, Bacchini claims that although many contextual properties have changed in the portion of Paris around Notre Dame since it was built, such change doesn't qualify as a context change, which is why we are not concerned by the fact that it has occurred. There are several alternative explanations that can support the view that the change isn't a context change: "either physical identity, or physical continuity, or in certain cases even perceptual indistinguishability or perceptual similarity seem perfectly sufficient to guarantee context identity over time", Bacchini observes (99; n. 24 p. 103), making reference to Derek Parfit's *Reasons and Persons* (1984: 207).

I agree with Bacchini's arguments in support of the view that all architectural objects are sited. Our views diverge, however, when it comes to identifying what kind of architectural objects qualifies as site-specific. Bacchini develops his proposal by applying a general view of site-specific art to the architectural case. However, his conception of site-specific art is inadequate, and this impacts negatively on his view of site-specific architecture. According to Bacchini, site-specific artworks are merely all and only those artworks that are constitutively located in a certain place (2017: 90); it follows that all buildings are akin to site-specific artworks. This claim, however, is debatable. As can be anticipated from my discussion in section (2), Bacchini fails to distinguish between sited artworks and site-specific artworks and, consequently, between architecture and site-specific architecture. As I shall argue, site-specific architectural objects are not just sited: they also articulate content about the historical and/or social and/or cultural context of their production through the sites that they encompass within their medium.

Let us first look at buildings simply qua sited. As we have seen in section (2), sited artworks like Rodin's *Les Bourgeois de Calais* do not articulate comments on the historical and/or social and/or cultural context of their production. Likewise, there are buildings that have clearly been designed to respond to their sites, but that don't articulate comments on their contexts of production: consider, for instance, the numerous cave houses and churches in Matera, Italy. All the buildings carved out from the rocks are appropriately described as responding to their location: since the Paleolithic period, human beings have responded to the presence of calcareous, friable rocks in Matera by carving rooms out of them, adapting to the topography of the territory. Some of the houses were inhabited until the 1950s. We have no evidence, however, that the excavated houses and churches were intended to articulate any content concerning their historical, social and cultural circumstances of production.

Now, one might ask whether buildings are capable of articulating content at all? More generally, recalling what I have said above about Lopes' discussion of media, one might ask whether buildings have media (i.e., whether they are partly constituted by materials manipulated to convey content and/or make salient some of their properties?). I think this question should be answered in the affirmative. Architecture, certainly, isn't customarily representational. Suzanne Langer (1953), Roger Scruton (1979), and Nelson Goodman (1985), however, argue that architectural objects can express or refer. Goodman, in particular, famously distinguishes between the less widespread sculpture-like modality of representation displayed by some buildings (1985: 644)—e.g., Jørn Utzon's Sidney Opera House, which looks like a giant sculpture of sailboats—and two more widespread meaning-conveying strategies adopted by the makers of buildings: exemplification and expression (645–647). According to Goodman, some buildings are designed to refer explicitly to some of the properties of their structures, which is why they *exemplify* those properties:

> A commonplace case [of exemplification] is a swatch of yellow plaid woolen serving as a sample. The swatch refers not to anything it pictures or describes or otherwise denotes but to its properties of being yellow, plaid, and woolen, or to the words 'yellow', 'plaid', and 'woolen' that denote it. But it does not so exemplify all its properties nor all labels applying to it—not for instance its size or shape. The lady who ordered dress material 'exactly like the sample' did not want it in two-inch-square pieces with zigzag edges.
>
> Exemplification is one of the major ways that architectural works mean. […] For instance, according to William H. Jordy, "the Dutch architect Gerrit Rietveld […] fragmented architecture into primal linear elements (columns, beams, and framing elements for openings) and planes (wall increments) in order to make visible the 'build' of the building". That is, the building is designed to refer explicitly to certain properties of its structure. In other buildings made of columns, beams, frames, and walls, the structure is not thus

> exemplified at all, serving only practical and perhaps also other symbolic functions.[10] (Goodman 1985: 645–646)

Moreover, Goodman argues, in some cases we attribute metaphorical properties to buildings. Goodman calls "expression" the exemplification of metaphorically possessed properties:

> Not all the properties (or labels) that a building refers to are among those it literally possesses (or that literally apply to it). The vault in the Vierzehnheilegen church [near Bamberg, Germany] is not literally eaten away; the spaces do not actually move; and their organization is not literally but rather metaphorically dynamic.[11] Again, although literally a building blows no brass and beats no drums, some buildings are aptly described as 'jazzy'. A building may express feelings it does not feel, ideas it cannot think or state, activities it cannot perform. (Goodman 1985: 646)

Furthermore, as Saul Fisher observes, buildings can have a "narrative character" since "the design of circulatory pathways allows architectural objects to communicate a sequence of events through the movement of visitors or inhabitants" (Fisher 2015: n. p.). Finally, I argue that *all* buildings have a medium: the reason is that all buildings make salient certain properties. Berlin's *Plattenbauten*, for instance, usually make salient their property of being symmetric structures and, as previously stressed, we cannot rule out that they also make salient properties such as being hit by light at dusk in a particular way at a specific Lichtenberg location. Likewise, Matera's cave houses make salient their property of having been carved out of rocks.

Let's now look at two examples of sited buildings that articulate content about their historical and/or social and/or cultural circumstances of production. On the one hand, there are buildings that, although sited, do not significantly recruit their sites to articulate content about their contexts of production; on the other hand, there are sited buildings whose sites are essential to articulating content about their contexts of production. An example of the former is the CCTV tower in Beijing, whose workings can be compared to those of, for example, the *Ara Pacis*, while an example of the latter is the Norwegian National Opera and Ballet, whose workings can be compared to those of, for example, *Hahn/Cock*.

As remarked above, The CCTV tower tends to attract the public around itself, because of its astonishing shape, thereby constituting an alternative to skyscrapers that, around Beijing, function as gated com-

[10] See William H. Jordy (1983), "Aedicular Modern: The Architecture of Michael Graves", *New Criterion*, 2.

[11] Goodman here refers to what Christian Norberg-Schulz writes about the church: "Over the crossing, where traditionally the centre of the church ought to be, the vault *is eaten away* by the four adjacent baldachins. The space defined by the groundplan is thereby *transposed* relative to space defined by the vault […] This *dynamic* and ambiguous system of main spaces is surrounded by a secondary, outer zone [etc.]" (Christian Norberg-Schulz, *Meaning in Western Architecture*, London 1975, p. 311, quoted in Goodman 1985: 646, my italics).

munities. Thus, it can be claimed that this, too, is a building that has been designed to respond to a peculiar feature of its surroundings, providing an alternative to prevalent modes of habitation in contemporary Beijing. Unlike the houses and churches in Matera's Sassi, however, the CCTV tower can be described as articulating content about its context, by means of exemplifying properties it metaphorically possesses. This is how Koolhaas describes the tower: "It looks different from every angle, no matter where you stand. Foreground and background are constantly shifting. We didn't create a single identity, but 400 identities. That was what we wanted: To create ambiguity and complexity, so as to escape the constraints of the explicit" (Koolhaas 2008: n. p.). The CCTV tower thus literally possesses the property of being irregularly shaped, so that it looks very different from different viewpoints. This property of the tower is made salient by its huge dimensions, which make it visible from a great variety of viewpoints. Moreover, Koolhaas invites us to think of the tower as of an object that doesn't have "a single identity, but 400 identities". The tower, however, doesn't literally possess 400 identities: first, it seems mistaken to claim that buildings literally have identities like persons do and, second, even if buildings had identities, it wouldn't make sense to claim that, literally, they could have more than one—just like it wouldn't make sense to claim that persons can, literally, have multiple identities. Rather, that of having 400 identities is a property that the tower metaphorically possesses, in virtue of exemplifying its literal property of looking very different from different viewpoints. The tower, then, expresses multiple-identity character. In so doing, I submit, the tower articulates content about the historical, social and cultural context of its production: manifesting its expressive character is a way for it to allude to the complexity of contemporary Beijing, which can also be described as ambiguous, complex and multifaceted. Indeed, the city straddles the 'developed' and 'developing' world as well as historical sites and hyper-modern districts, embodying Koolhaas' view of the "city of the future": "So the city of the future, and in fact even the city of today, constitutes not a whole but an archipelago of different enclaves, where ideological values could be installed in limited, strong, and specific places, but with no pretence at being universal" (Koolhaas 1995: 121). Finally, the building displays its expressive character thanks to the structure of the tower: it is the tower, then, that provides the means for articulating its context-related content, as opposed to particular aspects of the location recruited by the building. The workings of the CCTV tower, then, can be compared to those of the *Ara Pacis*, which, as we have seen, articulates content that is about the historical context of its production by depicting Emperor Augusts and his family officiating a procession, rather than by recruiting specific features of its physical setting. This doesn't mean that paying attention to some of the tower's relationships to its physical location isn't relevant to its appreciation: as mentioned above, the building's ability to

attract people around itself is better appreciated when seen against the background of the many high-rises and skyscrapers of contemporary Beijing that shut the public out of their premises completely.

Let us now consider the Norwegian National Opera and Ballet. Clearly, the building has been designed to respond to its site: its structure accommodates the presence of the fjord by gently sloping into it and its shape, as remarked above, neatly fits into the city's broader landscape. Furthermore, this is a building that articulates content about its social context of production. First, it suggests a narrative of coming-together through the articulation of its spaces, which provide ample room for free gatherings of people, outside and inside, independently of the performances taking place in the theatres hosted by the building. Thus, the opera house alludes to a view for contemporary Oslo: a city where people are invited to share public spaces, rather than merely gather in commercial areas or spend time mainly in private premises. Its architects have claimed: "The competition brief stated that the opera house should be of high architectural quality and should be monumental in its expression. One idea stood out as a legitimation of this monumentality: the concept of togetherness, joint ownership, easy and open access for all" (Snøhetta 2008: n. p.). Second, thanks to its multi-layered horizontal shape, the building appears to bridge the fjord to the west, and the Ekerberg hill to the east, when seen from the Akershus castle. It can be described, then, as merging different parts of the city. Now, the building doesn't literally merge the fjord and the Ekerberg hill: the properties it literally possesses are that of having such-and-such a shape and being positioned in such-and-such a location, while the property it metaphorically possesses, in virtue of possessing those literal properties, is that of merging different parts of the city. Since that is a salient property of the building, it is appropriate to claim that it is exemplified by it. The building, then, expresses the quality of bridging different parts of the city of Oslo. This is another strategy it employs to convey content about its social context: the opera house presents itself as a link between the richer West End and the poorer East End of Oslo, which are, traditionally, two economically and socially segregated areas of the city. Finally, for the present discussion it is important to stress that both the content-conveying strategies employed by the opera house I have described rely, crucially, on the recruitment of the building's site into its medium: on the one hand, for the opera house to convey a sense of open accessibility and the idea that it provides a natural space for people to gather, it is crucial that the building be seamlessly embedded into the fjord landscape; on the other hand, for the opera house to look like a bridge between the fjord and the Ekerberg hill, it is essential for it to be perceived in relation to those parts of the city of Oslo. I submit, then, that the Norwegian National Opera and Ballet works like site-specific works of art do: just like *Hahn/Cock*, for instance, it is an artefact that relies on its situated

character in order to convey content about the social context of its production. I conclude that, for this reason, the Norwegian National Opera and Ballet qualifies as a genuine work of site-specific architecture.

With the analysis of the Norwegian National Opera and Ballet, I have illustrated my view of site-specific architecture, which relies on my understanding of site-specific art. I submit that the claims illustrated through the above case-study can be generalized to at least some of the buildings that contemporary architecture practitioners describe as closely related to their sites.

To conclude, I have presented a view of site-specific architecture, modelled on my understanding of site-specific art. My proposal has two attractive features: it provides a unified account of site-specific art and architecture and, more importantly, it illuminates what some contemporary architects mean when they describe their practice as closely site-focussed.[12]

References:

Aldallal, E., AlWaer, H. and Bandyopadhyay, S. 2016. *Site and Composition. Design Strategies in Architecture and Urbanism*. London: Routledge.

Bacchini, F. 2017. "The Persistence of Buildings and The Context Problem." *Foootprint* 20: 17–36.

Barnett, L. 2013. "Katharina Fritsch on her Fourth Plinth cockerel sculpture: 'I didn't want to make fun—but I was invited.'" *The Guardian*, 24th July 2013, online edition: https://www.theguardian.com/artanddesign/2013/jul/24/katharina-fritsch-fourth-plinth-cockerel-sculpture (last access: 19th August 2020).

Coles, A. (ed.) 2000. *Site-Specificity: The Ethnographic Turn*. London: Black Dog Press.

Crimp, D. 1986. "Serra's Public Sculpture: Redefining Site-Specificity." In L. Rostenstock (ed.) *Richard Serra/Sculpture*. New York: The Museum of Modern Art: 41–57.

Crow, T. 1996. "Site-Specific Art: The Strong and The Weak." In *Modern Art in the Common Culture*. New Haven: Yale University Press.

de Muynck, B. 2004. "The end of the skyscraper as we know it: from CCG to CCTV." *Archis* 5. Republished online: https://movingcities.org/bertdemuynck/on-china/cctv_archis/ (last access: 19th August 2020).

Fisher, S. 2015. "Philosophy of Architecture." In E. N. Zalta (ed.). *The Stanford Encyclopedia of Philosophy*. Online publication: https://plato.stanford.edu/entries/architecture/#KinThiArcMak (last access: 19th August 2020).

Forbes, K. 2015. *Site Specific*. San Francisco: ORO Editions.

Foster, H. 1996. *The Return of the Real. The Avant-Garde at the End of the Century*. Cambridge: MIT Press.

Gaiger, J. 2009. "Dismantling the Frame: Site-Specific Art and Aesthetic Autonomy." *The British Journal of Aesthetics* 49: 43–58.

[12] I would like to thank two anonymous referees for their comments on a previous version of this paper.

Gibson, E. 2017. "Frank Lloyd Wright integrated architecture into nature at Fallingwater." *Dezeen* June 7th 2017. Online publication: https://www.dezeen.com/2017/06/07/fallingwater-frank-lloyd-wright-pennsylvania-house-usa-150th-birthday/ (last access 19th August 2020).

Gillgren, P. 2011. *Siting Federico Barocci and the Renaissance Aesthetic*. New York: Routledge.

Gillgren, P. 2017. *Siting Michelangelo. Spectatorship, Site Specificity and Soundscape*. Riga: Nordic Academic Press.

Gillgren, P. and Snickare, M. (eds.) 2012. *Performativity and Performance in Baroque Rome*. New York: Routledge.

Goodman, N. 1985. "How Buildings Mean." *Critical Inquiry* 11: 642–653.

Irvin, S. 2005. "The Artist's Sanction in Contemporary Art." *The Journal of Aesthetics and Art Criticism* 63: 315–326.

Kaye, N. 2000. *Site-Specific Art. Performance, Place and Documentation*. London: Routledge.

Koolhaas, R. 1995. "Atlanta, Paris, Singapore." *Lotus International* 84.

Koolhaas, R. 2004. *Content*. Köln: Taschen.

Koolhaas, R. 2008. "An Obsessive Compulsion towards the Spectacular." *Spiegel International* 18th July 2008. Online edition: https://www.spiegel.de/international/world/rem-koolhaas-an-obsessive-compulsion-towards-the-spectacular-a-566655.html (last access: 19th August 2020).

Kwon, M. 2002. *One Place After Another. Site-Specific Art and Locational Identity*. Cambridge: MIT Press.

Langer, S. 1953. *Feeling and Form. A Theory of Art*. London: Charles Scribner's Sons.

Lopes, D. M. 2014. *Beyond Art*. Oxford: Oxford University Press.

Marshall, D. and Weatherson, B. 2018. "Intrinsic vs. Extrinsic Properties." In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), URL = <https://plato.stanford.edu/archives/spr2018/entries/intrinsic-extrinsic/>.

Meyer, J. 2000. "The Functional Site; or The Transformation of Site-Specificity." In E. Suderburg (ed.). *Space, Site, Intervention. Situating Installation Art*. Minneapolis: University of Minnesota Press: 23–37.

Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford University Press.

PPG 2020. *The Colors of Fallingwater*. Online publication: https://www.ppgpaints.com/color/color-collections/fallingwater/the-colors-of-fallingwater (last access: 19th August 2020).

Scruton, R. 1979. *The Aesthetics of Architecture*. Princeton: Princeton University Press.

Sklair, L. 2006. "Iconic Architecture and Capitalist Globalization." *City: analysis of urban trends, culture, theory, policy, action* 10: 21–47.

Sklair, L. 2012. "Iconic Architecture in Globalizing Cities." *International Critical Thought* 2: 349–361.

Snøhetta 2008. "Oslo Opera House / Snøhetta." *ArchDaily*, 7th May 2008. Online publication: https://www.archdaily.com/440/oslo-opera-house-snohetta (last access: 19th August 2020).

Spens, M. (ed.) 2007. "Landscape Architecture: Site/Non-Site." *Architectural Design* 77.

Steinhauer, J. 2017. "How Mierle Laderman Ukeles Turned Maintenance Work into Art." *Hyperallergic* 10th February 2017. Online publication:

https://hyperallergic.com/355255/how-mierle-laderman-ukeles-turned-maintenance-work-into-art/ (last access: 19th August 2020).

Suderburg, E. (ed.) 2000. *Space, Site, Intervention. Situating Installation Art*. Minneapolis: University of Minnesota Press.

Williams, A. J. (ed.) 2011. *Nancy Holt: Sightlines*. Los Angeles: University of California Press.

Croatian Journal of Philosophy
Vol. XXI, No. 61, 2021
https://doi.org/10.52685/cjp.21.1.2
Received: August 1, 2020
Accepted: November 10, 2020

Aldo Rossi: "My Architecture Stands Mute and Cold"

ALEXANDRA DIAS FORTES*
IFILNOVA, Lisbon, Portugal

Aldo Rossi offers a captivating account of the relationship between human life and material forms. Rossi says that he came to "the great questions", and to his discovery of Ludwig Wittgenstein and Georg Trakl through Adolf Loos (Rossi 1982: 46). I will outline some connections between Loos, Trakl and Wittgenstein that might help us to grasp the way in which Rossi's assertive attitude concerning architecture gradually leans towards "forgetting architecture". (The goal is not to try and justify how they might have influenced Rossi; rather the aim is to try to understand Rossi's work with those connections as a backdrop; to outline a constellation of affinities.) The running thread being the internal relation between the object and the subject, i.e., "construction and the artist's own life" (Lombardo 2003: 97). I will conclude by considering architectural form on the page, that is to say, in Rossi's plans, "a graphic variation of the handwritten manuscript", and drawings, "where a line is no longer a line, but writing" (Rossi 1981: 6), and finally by considering what he says about his architecture, namely, that it stands "mute and cold," tough it will still "creak" (Rossi 1981: 44), and give rise to "new meanings".

Keywords: Aldo Rossi; architecture; analogy; artefacts; urban forms of life; aesthetics of urban everyday life.

1.

Aldo Rossi seems to inhabit a tension between his more affirmative, assertive view of the city as architecture and his later inclination to-

* Written with the support of the FCT (Fundação para a Ciência e Tecnologia [Foundation for Science and Technology]), under the PTDC/FER-FIL/32042/2017.

wards forgetting architecture,[1] marked by melancholy[2] (a feeling that he expresses in various ways). His explanation of fog's entering a building, for example, springs to mind as a powerful image of how an atmosphere enters a construction, saturating it with the feeling and character of the passage of time:[3] His drawings are also a reminder of how time acts upon things and of how things acquire undeniable distinctiveness through their permanence in time, even though they become worn—"like stones worn smooth by the feet and hands of generations of men"—surviving, in the end, as images or as similes and not so much as material presences.

This distinctiveness, or individuality, is conveyed by means of a new architectural viewpoint, a view in which size and perspective are remade and everyday objects occupy a space as huge as their presence in our lives—coffeepots, glasses, spoons, and bottles are as tall as buildings, as majestic as cathedrals. (See for instance the drawing *Dieses ist langer her—Ora questo è perduto*: 1975.) Their role, usually ignored because of their excessive familiarity, is no longer left unnoticed but is rather acknowledged; commonplace items assume their rightful place in the depiction of sensible forms with which human beings routinely interact, thanks to Rossi's capacity to find analogies between things and to place them together. We do not usually recognize their affinity. His collections—or recollections—help us to appreciate their similarities, however, especially insofar as they make us forget their function, directing our attention to their reality and subsistence in our life and

[1] As Patrizia Lombardo concisely puts it: "Rossi's poetic, almost intimate vision—architecture is like a love story, architecture should be forgotten—seems to be very much at odds with the positive tone of *L'architettura della città* (…). Aldo Rossi often remembers with tender irony his sound convictions at the time he wrote that book" (Lombardo 2003: 97).

[2] Melancholy is notoriously and internally associated to Rossi's work in Lopes' writings on the architect and his oeuvre (cf. Lopes 2016), and rightfully so, given the association of melancholy states with the speculative and the contemplative, words that could be used to describe the coloratura of Rossi's writings. In antiquity, owing to Hippocrates and his followers, health was conceived of as the result of the equilibrium of four humors: blood, yellow bile, phlegm, and black bile. In excess, black bile (located in the spleen, cold and dry like the earth and influenced by Saturn) caused fear and despondency, which, if persistent, were symptomatic of melancholia. [Melas (black) + khole (bile).] This contributed to the view that someone affected by melancholy was dark and gloomy: saturnine. According to Raymond Klibansky, Erwin Panofsky and Fritz Saxl (1964: 18–29), Aristotle was the first to recognize this mood in notable figures. Instead of afflicting these men, black bile worked together with their intellectual activity, causing their mental faculties to increase instead of declining. From then on, melancholy was associated with a capacity for speculative thought (cf. Saxl 1964: 168 ff.)

[3] "Just standing in *Sant' Andrea* at Mantua I had this first impression of the relation between *tempo*, in its double atmospheric and chronological sense, and architecture; I saw the fog enter the basilica, as I often love to watch it penetrate the *Galleria* in Milan: it is the unforeseen element that modifies and alters, like light and shadow, like stones worn smooth by the feet and hands of generations of men" (Rossi 1981: 1.)

in our memory *qua* pictures and depictions. Their features suddenly appear bare and as grand as the primal forms utilized by humans from time immemorial (or at least since we began making artefacts).

Rossi first suggests the idea that artefacts, and more importantly cities as urban artefacts, are vital to understanding how architecture evolves from appropriating a site to planning, construction, dwelling, and finally relinquishment (and indeed to understanding Rossi's own relation to architecture, what it meant to him and why) in his 1966 book, *The Architecture of the City*. I shall start by indicating a few of the attitudes and principles he expounds there—some of the concepts he delineates—which together act as a background for decoding his work, including its continuities and discontinuities over the years. I will briefly turn my attention to Adolf Loos to clarify how constitutive aspects of Rossi's practice and beliefs relate to his views on the Austrian architect. Following this, when considering his later understanding of architecture (for instance in his *Scientific Autobiography* of 1981), and bearing in mind that it was through Loos that he discovered both Georg Trakl and Ludwig Wittgenstein, I will concisely deal with the following questions: Which aspects of Georg Trakl's poetry did he find so appealing, and why? And how can Wittgenstein's philosophy, early and late, help us to grasp the development of Rossi's thought?[4]

2.

The main theme of *The Architecture of the City*, as the title aptly prefigures, is the standing and prominence of architecture to approach and to outline the city,[5] in one word, to understand it. Ultimately coinciding with architecture, the city is its finest, most refined presentation. Thus, the title also signals the process of its becoming, its extended making.[6] Architecture (and consequently, for Rossi, the architect) is what best provides a means of understanding the city (since architecture and the city eventually are one and the same), and of grasping it as a totality—as a changing, growing, expressive form of the collective memory of its dwellers and authors, of elements and places sought by the latter as a means of appropriating its various sites, especially by means of urban

[4] Loos and Wittgenstein met through Ludwig von Ficker, founder and editor of *Der Brenner* (cf. Monk 1991: 183.) Loos and Trakl also knew each other; proof of this is Trakl's dedication of *Sebastian im Traum* to Loos. Wittgenstein and Trakl were also acquainted.

[5] Cf. Lopes 2016: 135.

[6] "The city, which is the subject of this book, is to be understood here as architecture. By architecture I mean not only the visible image of the city and the sum of its different architectures, but architecture as construction, the construction of the city over time. I believe that this point of view, objectively speaking, constitutes the most comprehensive way of analyzing the city; it addresses the ultimate and definitive fact in the life of the collective, the creation of the environment in which it lives" (Rossi 1982: 21).

artefacts.[7] Architecture thus constitutes the first, crucial feature of the city that Rossi analyses, a feature that he explores by means of an example: The *Palazzo della Ragione* in Padua.[8] The wealth of functions that this building has served throughout its existence surprises Rossi. They show that form does not determine use, although form is precisely what affects and excites, what moves us, consequently configuring the city. This is why Rossi views it as a classic example of a phenomenon that he continually encounters and which ultimately leads him both to devalue the idea of function as being decisive to a building's history and to declare that functionalism is an incorrect method within architecture, unfit for dictating the design of urban artefacts. The latter do not follow function, although it could be said that function follows from urban artefacts insofar as these use form as needed in order to experience and *live* a space, shaping the city in the process. For this reason, an urban artefact possesses a specific identity, a uniqueness acquired throughout the process by means of which something that could otherwise be devoid of character becomes a specific place, imbued with spirit—the *genius loci*—in other words, according to Rossi, a *locus*.[9]

Although he first criticizes Adolf Loos, Rossi later praises him and his use of a Doric column in the project for the Chicago Tribune.[10] In particular, he celebrates it as an application of style that expresses a great understanding of the American city, which for Rossi is an example of the "architecture of the city" *par excellence*.[11]

[7] "[W]hich like the city itself are characterized by their own history and thus by their own form (…)" (Rossi 1982: 29.) In Rossi's text the English word 'artifact' is used to translate the Italian 'fatti', and 'urban artifact' stands for 'fatto urbano'. (This text uses instead 'artefact' and 'urban artefacts.) It should be noted that 'fatti', in Italian, refers not only to something made by human hands, but also to actions and facts, to deeds, to something that was accomplished; artefacts are related to the history of the city, they partake in its life. Hence, they are not merely physical objects.

[8] "In almost all European cities there are large palaces, building complexes, or agglomerations that constitute whole pieces of the city and whose function now is no longer the original one. When one visits a monument of this type, for example the *Palazzo della Ragione* in Padua, one is always surprised by a series of questions intimately associated with it. In particular, one is struck by the multiplicity of functions that a building of this type can contain over time and how these functions are entirely independent of the form. At the same time, it is precisely the form that impresses us; we live it and experience it, and in turn it structures the city" (Rossi 1982: 29.)

[9] Cf. Rossi 1982: 103 and 130.

[10] "Rossi's analysis evolves and changes radically through the various texts he wrote on Loos. In the first text for *Casabella*, Rossi seems to interpret the project for the Chicago Tribune as a reactionary contestation (…) Rossi reverses his position (…)" (Onaner 2012: 50.)

[11] "American architecture is above all 'the architecture of the city': primary elements, monuments, parts" (Rossi 1982: 15).

The importance of Loos's Doric column also lies in the fact that it perfectly illustrates the creation of an analogue.[12] Rossi was struck by it, especially by the ability of an ancient element to produce something significant and momentous in a new world, in a new time, an artefact—architecture—that was closely or internally connected with antiquity. Thus it is not a stretch to say that it deeply inspired him; in the "diverse totality" that makes up a city, he saw analogues such as the column everywhere he looked, leading him to develop the idea of the "analogous city."[13]

To better understand the impression that an artefact can make, an impression that affirms its identity as architecture and as an analogue—or in other words, a recuring form, built, depurated, through use and memory, into a precise construction that persists and returns—we would do well to bear in mind Loos's remark on what might drive us to say, "That is architecture."[14] Likewise, we should remember Rossi's comment on this remark. He calls our attention to the fact that, owing to a building's extreme architectural formal purity, we are able on the one hand to identify it—to see the artefact in it—and on the other hand to repurpose it for understanding our own experience. We can measure the latter against something that has resisted time and has become a paradigm of sorts, a classic.[15] Hence, we can say that artefacts possess

[12] The use of the Doric column in the project for the Chicago Tribune exemplifies the creation of an analogue because it does not simply copy a form and applies it there. It shows recognition of what in history is still alive and revives it in a new place and time. Creating an analogue—making an analogy—is different from copying because it entails establishing a relationship by way of an abstract move that filters memory and finds a proportion, a meaningful correspondence between things and their meaning. Rossi declares that he found in Carl Jung's definition of analogical thought as "a meditation on themes of the past", "a different sense of history conceived of not simply as fact, but rather as a series of things, of affective objects to be used by the memory or in a design" (Rossi 1976: 75).

[13] For example, as a diverse totality: "Clearly, the Cathedrals of Milan and Reggio Emilia and the *Tempio Malatestiano* in Rimini were—are—beautiful in their incompleteness. They were and are a kind of abandoned architecture—abandoned by time, by chance, or by the destiny of the city. The city in its growth is defined by its artifacts, leaving open many possibilities and containing unexplored potential. This has nothing to do with the concept of open form or open work; rather it suggests the idea of interrupted work. The analogous city is in essence the city in its diverse totality; this fact is visible in the echoes of the East and the North that one finds in Venice, in the piecemeal structure of New York, and in the memories and analogies that every city always offers" (Rossi 1982: 18.) For other aspects of the analogous city, cf. Rossi 1982: 13, 15, 18, 20, 22–23, 71, 92, 124, 136, 143, 161, 166, 174 and 176.

[14] "If we find a mound six feet long and three feet wide in the forest, formed into a pyramid, shaped by a shovel, we become serious and something in us says, 'someone lies buried here.' That is architecture" (Adolf Loos, quoted by Rossi 1982: 107).

[15] "The mound six feet long and three feet wide is an extremely intense and pure architecture precisely because it is identifiable in the artifact. It is only in the history of architecture that a separation between the original element and its various forms occurred. From this separation, which the ancient world seemingly resolved forever, derives the universally acknowledged character of permanence of those first forms.

methodological value: we can use them to investigate and make sense of the present. By way of comparison, specifically of "different places" and "our own cities," we can begin to understand our "individual experience of each particular place," meaning that we can find an analogue in artefacts that helps us to come to grips with our own way of living. This process of finding a useful correspondence between different things is made possible by a practice that Rossi describes as consisting in the repeated use of the architecture of antiquity, an era that Rossi considers to be exceptional in its remarkable dedication to construction and the elucidation of form. Furthermore, the practice of finding an equivalence that can clarify the actuality of the singular experience of a specific place can also be a basis for constructing anew, that is to say, for creating a new *locus* ensuing from the likeness, the similarity, and the memory contained in the many recurring and replicated artefacts.[16]

The *locus* deserves further attention: although it arises from repeating ancient formulas, these are recast, reorganized, and reformed to bring about something novel. It is a notion bound up not only with repetition but with disruption (and where would we be without it?), with breaking the mold, meaning that, although it appropriates something that was previously available, it simultaneously establishes a difference and a particularity. In addition, it is a great image for understanding Rossi's fascination with the unforeseeable event that architecture should enable and facilitate, even if it means forgetting—as we shall see—its own constructs so that the latter might finally disclose the singular, individual existence that specifically makes them artefacts. Once again—as with the *Palazzo* in Padua—there is something in their form that was unpredictable in light of the function for which they were built.

All of the great eras of architecture have reproposed the architecture of antiquity anew, as if it were a paradigm established forever; but each time it has been reproposed differently. Because this same idea of architecture has been manifested in different places, we can understand our own cities by measuring this standard against the actuality of the individual experience of each particular place. What I said at the beginning about the *Palazzo della Ragione* in Padua is perhaps subsumed in this idea, which goes beyond a building's functions and its history, but not beyond the particularity of the place in which it exists" (Rossi 1982: 107).

[16] "Perhaps we can better understand the concept of *locus*, which at times seems rather opaque, by approaching it from another perspective (…) Otherwise, we continue to grasp at outlines which only evaporate and disappear. These outlines (…) trace the relation of architecture to its location—the place of art—and thereby its connections to, and the precise articulation of, the *locus* itself as a singular artifact determined by its space and time, by its topographical dimensions and its form, by its being the seat of a succession of ancient and recent events, by its memory" (Rossi 1982: 107).

3.

All of these ideas—the city as architecture, urban artefacts, analogy, and the *locus*—are also found in Rossi's later views, although they undergo variations. Things are likewise with his rejection of the reduction of form to function, given that its potential is revealed through varied actualities and uses and not through a single (pre-determined) purpose. The latter only enters the equation obliquely, by way of repurposing, such as the repurposing of an artefact on the basis of its form. (For this reason, a building can be built as a convent and later be used as a school.)[17]

Each of those variations relates to a central motif, namely abandonment ("I could have called my book *Forgetting Architecture*").[18] In reality, they each relate to a forgetting entwined with remembrance, a relinquishment of Rossi's earlier conception of architecture and its reconfiguring as "ritual," comprising repetitions and re-enactments that guarantee "continuity, … allowing us to live with every change which, because of its inability to evolve, constitutes a destruction."[19]

This conception of architecture as ritual likely results from the awareness that architecture as a kind of creation leaves a place after making it into one, or, to be more exact, after the place exhausts its possibilities for growth (the realization of "all that was" lies behind this gloomy prospect). After preparation, appropriation, life, and the workings of time, we become oblivious to architecture as construction while still knowing and holding on to its physiognomy, that is to say, its character, its appearance, its look. We recognize it throughout its modifications because in our memory we retain and assert its identity. (And the way we realize this is through everyday rituals, performed, routinely, and driving us to recognize someplace as *a* place—where we act out our lives; where gestures become imbued with meaning.) If we were to put this in terms of continuity and discontinuity, i.e., in terms of what remains of Rossi's former view of architecture in his new understanding and relation to it, we could say that architecture still matters very much; in fact, it acquires the stately character of ritual, and while it leaves something behind, including a more conventional understanding of built form, it puts something new in its place in tandem

[17] "I have seen old palaces now inhabited by many families, convents transformed into schools, amphitheatres transformed into football fields; and such transformations have always come about most effectively where neither an architect nor some shrewd administrator have intervened" (Rossi 1981: 44)

[18] "As I have said, *Forgetting Architecture* comes to mind as a more appropriate title for this book, since while I may talk about a school, a cemetery, a theatre, it is more correct to say that I talk about life, death, imagination" (Rossi 1981: 78).

[19] "Today if I were to talk about architecture, I would say that it is a ritual rather than a creative process. I say this fully understanding the bitterness and the comfort of the ritual. Rituals give us the comfort of continuity, of repetition, compelling us to an oblique forgetfulness, allowing us to live with every change which, because of its inability to evolve, constitutes a destruction" (Rossi 1981: 37).

with the formula "architecture as ritual," a notion of "building" that includes drawings, collages, and writings:[20] architecture on the page. It does this in moments of preparation[21] but also in moments of contemplation, profoundly marked by melancholy, a feeling that enhances its appeal, since it instils it with a mysterious and meditative character that observes, absorbs and reckons with all that is lost, viz. by means of salvaging what lies destroyed in drawings, collages, and writings, including the autobiographical. Moreover, this makes architecture all the more conscious of its circumstances and the opportunity that these offer for the occurrence of something new—of all that is still promising about it now that the possibility of great things has been "precluded"[22] and, moreover, since the element of destruction has entered the scene. Hence, not all is lost: destruction leaves the space open for the novel to emerge, and for what is left unsaid, so that it may peer through time and the fractures, rise from the ruins.

This is the theme of many of Rossi's drawings, including a work that borrows its title from Trakl's *Abendlied*, *Dieses ist lange her* (1975), but also, e.g., *L'architettura assassinata* (1976).

Instead of longing for architecture and for his positive attitude towards times past, Rossi engages in a series of descriptions that bring together diverse elements in a meaningful way, providing new acumen by making correspondences between things otherwise dissimilar. Where the author of *The Architecture of the City* firmly believed in his craft's ability to promote urbanization, the author of the *Scientific Autobiography* believes in our ability to understand this phenomenon by means of analogy (preserving some of his past optimism). To be sure, analogy makes an appearance in his 1966 book, and an important one at that; almost twenty years later, however, things appear to have changed.[23] No longer "only" a tool for reading the city, analogy also creates it virtually in plans ("a graphic variation of the handwritten manuscript") drawings ("where a line is no longer a line, but writing")

[20] Cf. Eisenman in Rossi 1982: 10.

[21] "The event might not ever happen anyway. I am more interested in the preparations, in what might happen on a midsummer night. In this way, architecture can be beautiful before it is used; there is beauty in the wait, in the room prepared for the wedding, in the flowers and the silver before High Mass" (Rossi 1981: 65–66).

[22] "To what, then, could I have aspired in my craft? Certainly to small things, having seen that the possibility of great ones was historically precluded" (Rossi 1981: 23).

[23] See for instance the following excerpt from Peter Einsenman's Introduction to *The Architecture of the City*: "The bourgeois house of Rossi's childhood permitted fantasy, but denied the ordering of type. *The Architecture of the City* attempts, through the apparatus of type, to place the city before us in such a way that, in spite of history, memory can imagine and reconstruct a future time of fantasy. This memory is set into motion through the inventive potential of the typological apparatus, the analogous design process of Rossi's drawings of the 'analogous city' can be seen to evolve directly from his writing of *The Architecture of the City*" (Einsenman in Rossi 1982: 10).

and on the written page of a manuscript ("between writing and drawing") (Rossi 1981: 6). But how? What makes "drawing out the *locus*" possible? (Dodds 1992: 5).[24]

4.

To begin answering this question, we must discuss Georg Trakl's significance to Aldo Rossi, from whom he seizes a line to use as a title to one of his most famous drawings, *Dieses ist Lange her*. Quoting is an analogical move of sorts, the quote being an analogue that immediately saturates the depiction with the atmosphere of Trakl's poetry. If we wish to describe this atmosphere, we should pay attention to the fact that it is filled with a deliberate silence. As Robert Bly has pointed out, the poet's voice is somewhat absent, and the things portrayed are themselves still and quiet, creating a formidable and meditative ambiance.[25]

It is indeed remarkable that poetry can be a silent affair, capable of quieting all that usually conceals things from our eyes. Trakl's restraint with words renders visible what would otherwise be hidden; he makes space for everything that he does not put into words, such that "nature has more and more confidence in him" and "more and more creatures live in his poems" (Bly 1961: 2). Silence facilitates attention and awe; things are described, although the poet says very little about them. The atmosphere is not exactly sad (things "live in too deep a joy to be gay"); it is infused with a pensive mode that we might call melancholy. It is striking that even in his darkest poems, which provide a precise image of a destroyed landscape, we are filled not so much with dread as with a calm, contemplative sorrow imbued with reverence: for the ruins (of war, for instance), for all that is lost and that, in a sense, we can make out in the poems.

[24] "[T]he analogous design process of Rossi's drawings of the 'analogous city' can be seen to evolve directly from his writing of *The Architecture of the City*. The analogous drawing embodies a changed condition of representation; it exists as the record of its own history. Thus, Rossi's drawings of the city, giving form to their own history, become part of the city, not just a representation of it. They have an authenticity, a reality which is, precisely, that of illusion. This reality may then, in turn, be represented in actual buildings" (Einsenman, "Introduction" in Rossi 1982: 10.)

[25] "The poems of Georg Trakl have a magnificent silence in them. It is very rare that he himself talks—for the most part he allows the images to speak for him. Most of the images, anyway, are images of silent things. In a good poem made by Trakl images follow one another in a way that is somehow stately. The images (…) live in too deep a joy to be gay. (…) Everywhere there is the suggestion of this dark silence … The silence is the silence of things that could speak, but choose not to. The German language has a word for deliberately keeping silence, which English does not have. Trakl uses this word *schweigen* often. When he says *the flowers/ Bend without words over the blue pond*, we realise that the flowers have a voice, and that Trakl hears it. They keep their silence in the poems. Since he doesn't put false speeches into the mouths of plants, nature has more and more confidence in him" (Bly 1961: 2).

Another characteristic of Trakl's poetry that is relevant to the matter at hand, namely his connection with Rossi, is the way in which he presents things. Images follow one another demurely, unhurriedly, and fluidly—a very surprising effect, given his profuse use of parataxis. It is as if all the disconnected parts were held together by their muteness. Without creating the impression of being discombobulated, and although they seem severed from one another, his images do not undermine the unity of his poems, the best example of this being *Grodek*—his great and last poem[26] (on war).

What Rossi might have found so appealing about Trakl's poetry is that he shares with the latter an ability to portray things—even ruins—such that they are instantly presented in their connectedness without losing any of their particularities. We see affinities, but these do not erase the traits that distinguish each. Moreover, the cohesiveness of Rossi's images (his drawings and collages) and of Trakl's poems allows an eloquent, meaningful feeling to shine through the cracks of their silence.[27] This cohesiveness provides an idea of the whole, even when both deal with ruins (a trait that is perhaps related to the Romantic aspect of their way of looking at things).[28] This last aspect—the possibility that silence expresses something, be it in a poem, a drawing, a collage, a plan or a building—is decisive for grasping the way in which analogy functions as a method for Rossi, one which allows him to request that we turn our attention to his work rather than asking him for further words of explanation. If we exert this type of thinking (comparative, analogical), we too will be able to see the parallels between fog and the character of a building, for example—as Rossi was able to do in the passage cited at the beginning of this paper—and to learn from the unpredictable, revealed by similarity.

[26] "There is here no foreground or background but a single poetic landscape with colours that have the sharpness of one of El Greco's pictures of Christ at Toledo. The arrangement is almost entirely paratactic: simple subject-predicate phrases are organised into large syntactic units, but these units appear to be disconnected from each other. There is, in the poem's seventeen lines, not a single causal or temporal conjunction; there are two pronominal adverbs which establish an un- emphatic spatial relationship—'*darüber*', '*darin*' ('above which', 'in which'), and an infinitive expressing purpose—'*zu grüßen*' ('[in order] to greet'): the only connections between the parts of speech given in the poem are simple copulas and the punctuation; it is as though its meaningful totality were formed by a pattern of relatively independent images. The poem achieves its unity without reference to an authorial self" (Stern 1995: 253).

[27] Since Rossi, too, believed that his work achieved a "degree of silence": "I have tried to propose buildings which, so to speak, are vehicles for events (…) I can say now that they achieve a silence, a degree of silence which is different from the purism I had striven for in my early designs, where I was concerned primarily with light, walls, shadows, openings. I have realized that it is impossible to recreate an atmosphere. Things are better experienced and then abandoned; initially, everything should be foreseeable, even though what is not foreseeable is all the more fascinating because it remains beyond us" (Rossi 1981: 5).

[28] For this reason, it is not so surprising that Aldo Rossi uses lines from Hölderlin when trying to sum up his architecture.

5.

Ludwig Wittgenstein once said of a collection of Trakl's poems: "I do not understand them, but their *tone* makes me happy. It is the tone of true genius."[29] For Wittgenstein, this probably meant an acceptance of how things are. At the time, he conceived of the world as being independent of our will,[30] the only way to conquer it consisting in recognizing and conforming to it by adopting a happy outlook,[31] one that removes the *Sorge* not so much by trying to alter the form of life that belongs to the world (which is impossible) but by changing how we live. He might have found a spirit akin to his own in Trakl. Trakl is likewise happy to simply put everything before our eyes without musing on the state of affairs. In both figures, we find the necessary means for attaining a view of the world *sub specie aeterni*, a view that sees things in terms of their infinite possibility—or, as Wittgenstein would say, with the whole space of possibility as their background[32]—in the grand scheme of things, in light of which their present configuration becomes meaningful and part of something that surpasses it.

Wittgenstein thought that art provided the right perspective, compelling us to look at objects correctly,[33] and in 1939 he declared that he had come to think that there was another way, a philosophical way, to attain a correct view of the world.[34] He introduces it after pondering something that his friend Paul Engelmann had told him, something that made him realize how magnificent it would be if we could observe people engaged in commonplace actions, unaware that they were being watched. He thought that such a spectacle would be superior to anything we could experience by going to the theatre—but it also made him realize that we see these things every day, although not in the same way, that is, as valuable (worthy of being put on stage). It was this that called for a reconfiguration of the customary *Weltanschauung*. He would later refer to this new way of attaining a correct view of the world as "panoramic presentation" (*übersichtliche Darstellung*), namely in § 122 of the *Philosophical Investigations* (Wittgenstein: 2009), an observation that he revised countless times over the years (from the beginning of the 1930s). In the interest of keeping with our central subject—without getting into much detail, but also without completely foregoing a delineation—we can say that this concept of an overarching view is of fundamental importance; it denotes on the one hand a clear vision (for example of a landscape, a language game, or a city) and on the other hand a method that proceeds by analogy, which compares

[29] Quoted in Monk 1991: 186.

[30] Cf. Wittgenstein 2003: proposition no. 6.373.

[31] Cf. Wittgenstein 1979: entry dated 11.6.1916 and Wittgenstein 2003: proposition no. 6.43.

[32] Cf. Wittgenstein 1979: entry dated 7.10.1916.

[33] Cf. Wittgenstein 1979: entry dated 20.10.1916.

[34] Cf. Wittgenstein 2006: 6–7.

and brings together words, images and fragments, coalescing and reconfiguring them by way of intermediate links in a meaningful whole (or indeed several meaningful wholes). The aim is to provide not a final synopsis of things—which would probably be impossible—but rather many synopses that elucidate our form(s) of life, including urban ways of life. Its significance for what occupies us now has to do with the fact that it is very close to what Rossi does when, by way of an analogical procedure, he makes us see the interconnectedness of the things he joins together, juxtaposing them precisely as a Wittgensteinian synopsis of facts is supposed to do. It is also worth noting that both Rossi and Wittgenstein receive some measure of inspiration from discovering the connection between theatre and life. For Rossi, this connection implies the possibility of transformation and change in life by way of the theatre: "the theatre, and perhaps only the theatre, possesses the unique magical ability to transform every situation." Architecture is also a conduit for change, however, as well as for "everything that is unforeseeable in life."[35]

6.

The "unforeseeable" aspect finally brings us back to what we said above about architecture's being possible on the page. Analogy plays a role: finding analogies is what might assist us to see on the page a simile of what we discover through our experience of the city, comprising the built forms with which we interact in our urban form of life (houses, monuments, cenotaphs, gardens…). Structures that we can enter represent for Rossi, in his drawings of small utensils (which are also built forms, but which we cannot enter), the analogue of the things that have always interested him and that belong both to his experience of the interior and to his interior experience, his personal relation to architecture. Indeed, he sees these small utensils as "miniatures of the fantastic architectures that [he] would encounter later."[36]

Similarly, in his work we can find the necessary analogues for understanding architecture, the city, and our lives as a part of the latter.

Over time for Rossi, architecture becomes propitiatory, something

[35] "[A]rchitecture [like the theatre] becomes the vehicle for an event we desire, whether or not it actually occurs; and in our desiring it, the event becomes something 'progressive' … it is for this reason that the dimensions of a table or a house are very important—not, as the functionalists thought, because they carry out a determined function, but because they permit other functions. Finally, because they permit everything that is unforeseeable in life" (Rossi 1981: 3).

[36] "I have always had a strong interest in objects, instruments, apparatus, tools. Without intending to I used to linger for hours in the large kitchen (…) drawing the coffeepots, the pans, the bottles. I particularly loved the strange shapes of the coffeepots enamelled blue, green, red; they were miniatures of the fantastic architectures that I would encounter later. Today I still love to draw these large coffeepots, which I liken to brick walls, and which I think of as structures that can be entered" (Rossi 1981: 2).

that can trigger what we wish for, but also what we did not know we wanted and were not expecting: the unpredictable appearance of the new, full of meaning. This warrants a quotation, a sort of final word from Rossi on the matter of his architecture that tells us to look at his work and to discover what it meant for him: the impressive declaration "My architecture stands mute and cold."[37]

Aldo Rossi's recognition of all that was lost—and that did not prevent him from acknowledging what was still possible, what he could still hope for for his architecture—is related to something Wittgenstein said about architecture: that it was only possible in times when there was something that could be expressed, that is to say, that merited an expression.[38] We might say that Wittgenstein also viewed such a possibility as belonging to the past. All of this does not mean that we must rest satisfied or happy; it only means that we should not hastily build (in the conventional sense), potentially disfiguring what already exists by introducing a new construction, without first noticing, and letting sink in, the place's spirit, the memory that inhabits its (worn out) form, not allowing it to survive. Through its difference and singularity, each place tells a story—one that, if obliterated by a larger context, is lost in an amorphous whole. Where this singularity can shine through, accumulation, layers, and strata can form the ground for future possibilities that stem from its life.

References

Dodds, G. 1992. "Drawing from Memory: Context, Locus, and the Difficult Journey." *History / Post-History: Proceedings of the ACSA Southeast Regional Conference*. Savannah: Savannah College of Art and Design (unpaginated).

Klibansky, R., Panofsky, E. and Saxl, F. 1979. *Saturn and Melancholy: Studies in the History of Natural Philosophy*. Nendeln/Lichstenstein: Kraus Reprint.

Lopes, D. S. 2016. *Melancolia e Arquitectura em Aldo Rossi*. Lisbon: Orfeu Negro.

Lombardo, P. 2003. *Cities, Words and Images: From Poe to Scorcese*. Basingstoke: Palgrave Macmillan.

[37] ""To creak' is the translation of the German *klirren*, which has always struck me in Hölderlin's poem *Halfte des Lebens*. The very title of the poem seems to me a condition of suspension. The little iron banners which Hölderlin never drew himself subsequently invaded my drawings, and I am unable to answer any further the persistent questions I am asked about them except to say that I have translated the last lines of Hölderlin's poem into my architecture: 'The walls stand / mute and cold, in the wind / the banners creak].' I concluded one of my lectures at Zurich with this quotation, which I applied to my projects: My architecture stands mute and cold" (Rossi 1981: 44).

[38] "Architecture glorifies something (because it endures). Hence there can be no architecture where there is nothing to glorify" (Wittgenstein 2006: 74).

Monk, R. 1991. *Wittgenstein: The Duty of Genius*. London: Vintage Books.

Onaner, C. 2012. "Rossi Reader of Loos: Architecture in Suspense." *Adolf Loos: Our Contemporary/ Unser Zeitgenosse/ Nosso Contemporâneo*. Yehuda E. Safran (exhibition curator), New York: GSAPP Columbia University: 49–53.

Rossi, A. 1976. "An Analogical Architecture." *Architecture and Urbanism* 56 (May 1976): 74–76.

Rossi, A. 1981. *A Scientific Autobiography*. D. Ghirardo and J. Ockman (transl.), P. Eisenman (intro.). Cambridge: MIT Press.

Rossi, A. 1982. *The Architecture of the City*. D. Ghirardo and J. Ockman (transl.), P. Eisenman (intro., rev. for the American edition by A. Rossi and P. Eisenman), Cambridge: MIT Press.

Stern, J. P. 1995. *The Dear Purchase. A Theme in German Modernism*. UK: Cambridge University Press.

Trakl, G. 1961. *Twenty Poems of Georg Trakl*. James Wright and Robert Bly (transl. and selection). Madison: The Sixties Press.

Wittgenstein, L. 1979. *Notebooks: 1914–1916*. New York and Evanston: Harper & Row Publishers.

Wittgenstein, L. 2003. *Logisch-philosophische Abhandlung*. Frankfurt am Main: Suhrkamp Verlag.

Wittgenstein, L. 2009. *Philosophische Untersuchungen / Philosophical Investigations*. G. E. M. Anscombe, P. M. S. Hacker and J. Schulte (transl.), P. M. S. Hacker and J. Schulte (rev. fourth ed.), UK: Willey-Blackwell.

Wittgenstein, L. 2006. *Vermischte Bemerkungen: Eine Auswahl aus dem Nachlaß / Culture & Value: A Selection from the Posthumous Remains*. Oxford: Blackwell Publishing.

Croatian Journal of Philosophy
Vol. XXI, No. 61, 2021
https://doi.org/10.52685/cjp.21.1.3
Received: July 21, 2020
Accepted: November 3, 2020

Art Addressing Consumerism in the Age of Late Capitalism

POLONA TRATNIK
Faculty for Slovene and International Studies, New University, Ljubljana, Slovenia

The globalized world is still in the phase of late capitalism, signified by the establishment of multinational corporations, globalized markets and work, mass consumerism and the fluid flow of capital. The question of the criticism of art towards the capitalist system, its ideology and consumerism is therefore still current and is readdressed in this contribution. Considering this issue, the recurrent theoretical reference is American materialist aesthetician Fredric Jameson, who was among the first to define culture and art in the context of late capitalism. In the article the author revises Jameson's critique of art addressing consumerism and demonstrates that he did not consider the relevance of the means of consumption as regards the cultural logic of late capitalism. She claims that in order to open space to examine contemporary art as being critical towards consumerism, one also needs to consider the ontological changes that have occurred to art and pay attention to performative art, while Jameson was still focused on a representational mode of art. By being performative and also setting out actions outside of spaces that were traditionally designed for art, in the space meant for consumption, art has much a better chance to act politically, which Jameson wished to see from art which addresses consumerism, but did not. The author argues that if one is to seek critical or political art in late capitalism, those would be the cases of artistic interventions into the means of consumption.

Keywords: Late capitalism; contemporary art; means of consumption.

The globalized world is still in the phase of late capitalism as defined by Ernst Mandel, signified by the establishment of multinational corporations, globalized markets and work, mass consumerism and the

fluid flow of capital, which has taken place from the 1960s onwards (Mandel 1972). The question of art's criticism towards the capitalist system, its ideology and consumerism is therefore still current. Considering this issue, the recurrent theoretical reference is American materialist aesthetician Fredric Jameson, who was among the first to define culture and art in the context of late capitalism. His theory on postmodernism has gained such theoretical relevance that it has not yet been replaced by any other more current theory of similar or comparable importance. I will therefore revise Jameson's critique of art which addresses consumerism. The deconstruction of this influential theory is important in order to open space to examine the functioning of the means of consumption, to borrow the term from Georg Ritzer, and the critical potential of artistic intervention into these means. In order to do that, we need to consider the ontological changes through which art has gone and pay attention to performative art, while Jameson, on the other hand, was still focused on a representational mode of art. By being performative and also by setting out actions outside of spaces traditionally designed for art, one can assert that art has a much better chance of engaging in political action in the public space.

1. *Art as the Unconcealment of the Truth*

According to the observation of the renowned Fredric Jameson, popart as art that appears in the phase of late capitalism reflects commodity fetishism but fails to develop criticism. He recognizes

> one of the central issues about postmodernism itself and its possible political dimensions: Andy Warhol's work in fact turns centrally around commodification, and the great billboard images of the Coca-Cola bottle or the Campbell's soup can, which explicitly foreground the commodity fetishism of a transition to late capital, ought to be powerful and critical political statements. If they are not that, then one would surely want to know why, and one would want to begin to wonder a little more seriously about the possibilities of political or critical art in the postmodern period of late capital. (Jameson 1991: 9)

In his observation of consumerism, Jameson examines the problem of commodification. Objects are transformed into commodities and are fetishized. This holds true not only for Coca-Cola, but for human subjects as well. Stars such as Marilyn Monroe are transformed into images of themselves. Jameson is focused on the works of art as artifacts. He does not, therefore, for instance, acknowledge that artists such as Andy Warhol, whose work he discussed, have themselves become fetishized commodities. Instead of broadening his perspective on commodity fetishism, he approached the works of art as if they were still the "consecrated" objects meant for contemplation. Because Warhol's paintings obviously don't have this function, he sees them as works of art which are hollow and poor. This perspective is supported by his recognition that a feature of postmodernism is a new depthlessness. He acknowl-

edges "intensities", which are free-floating and euphoric and which are in accordance with the spirit of consumerism. In contrast, modernism praised depth, living context, space for the observer who had a possibility to supplement the work of art. Loneliness, social fragmentation and isolation were expressed and the world of anxiety and alienation was revealed. Instead of deep expressions and existential themes, postmodernism shows and repeats the fetishism of consumer goods and its ubiquity. The situation is comparable to an addiction, "with a whole historically original consumers' appetite for a world transformed into sheer images of itself and for pseudoevents and 'spectacles'" (Jameson 1991: 18). Such objects are "simulacrum", that is "the identical copy for which no original has ever existed" (Jameson 1991: 18). Parody and speaking in a dead language are significant for postmodernism. Modernism, on the other hand, is characterized by originality and the search for truth. Jameson refers to Martin Heidegger and his study of Van Gogh's painting *A Pair of Shoes* (1886) to outline the characteristics of postmodernism (in art) by comparing Van Gogh's and Warhol's shoes.

For Heidegger art is the "setting-into-work of truth" (Heidegger 1993: 199). He studies the case of the painted peasant shoes and ascertains: "Van Gogh's painting is the disclosure of what the equipment, the pair of peasant shoes, *is* in truth." (Heidegger 1993: 161) The truth as addressed by Heidegger means "the unconcealment of beings" which the Greek called *alētheia* (Heidegger 1993: 161). "If there occurs in the work a disclosure of a particular being, disclosing what and how it is, then there is here an occurring, a happening of the truth at work" (Heidegger 1993: 161–162) in the work of art. Heidegger deliberately chose a commonly found equipment—a pair of peasant's shoes, which is a mere thing, a self-contained thing like a block of granite is a material in a definite form. Then he took Van Gogh's painting of the pair of peasant's shoes to articulate the significance of the work of art representing the pair of shoes:

> From the dark opening of the worn inside of the toilsome tread of the worker stares forth. In the stiffly rugged heaviness of the shoes there is the accumulated tenacity of her slow trudge through the far-spreading and ever-uniform furrows of the field swept by raw wind. On the leather lie the dampness and richness of the soil. /…/ This equipment is pervaded by uncomplaining worry as to the certainty of bread. (Heidegger 1993: 159)

There is a meaningful difference between the mere thing and the work of art. Everything that the peasant's shoes are is to be found in the painting. "The peasant woman, on the other hand, simply wears them." (Heidegger 1993: 160) Art is thus understood essentially as the agent unconcealing the truth. The work of art *presents* it. At the same time, what art is "should be inferable from the work." (Heidegger 1993: 144) There is work invested in the work of art that makes it a work. Heidegger sees art as essentially *artifactual*. In addition, this thingly element, as Heidegger called it, actually refers to the sensual character of art. "The thingly element is so irremovably present in the artwork that

we are compelled rather to say conversely that the architectural work is in stone, the carving is in wood, the painting in color, the linguistic work in speech, the musical composition in sound." (Heidegger 1993: 145) At this point we recall Georg W. F. Hegel's consideration that "art presents itself to *sense*, feeling, intuition, imagination" (Hegel 1975: 5). The beautiful has its being in "pure appearance" (Hegel 1975: 4). "But appearance itself is essential to essence. Truth would not be truth if it did not show itself and appear" (Hegel 1975: 8). Hegel required the freedom of art. Art in general can serve other ends and in this sense is a mere passing amusement. This required autonomy of art, its freedom, makes fine art "truly art, and it only fulfils its supreme task when it has placed itself in the same sphere as religion and philosophy, and when it is simply one way of bringing to our minds and expressing the *Divine*, the deepest interests of mankind, and the most comprehensive truths of the spirit." (Hegel 1975: 7) For Hegel, art assures a special kind of appearance through which it communicates what is "inherently true" (Hegel 1975: 8). Art is a special kind of representation for both Hegel as well as Heidegger. For Hegel, even if it subordinates itself to serious aims and produces serious effects, the means that art uses for this purpose is *deception* (Hegel 1975: 4). The "representations of art" are "a deceptive appearance"; truer representations would be of historiography, however, the content of the latter

> remains burdened with the entire contingency of ordinary life and its events, complications, and individualities, whereat the work of art brings before us the eternal powers that govern history without this appendage of the immediate sensuous present and its unstable appearance. (Hegel 1975: 9)

For Heidegger, truth is the truth of Being. "Beauty does not occur apart from this truth. When truth sets itself into the work, it appears. Appearance—as this kind of being of truth in the work and as work—is beauty. Thus the beautiful belongs to truth's propriative event." (Heidegger 1993: 206) In Heidegger's view, art has such high status that the truth can or even must happen as art. Yet, he concludes the discussion on the origin of the work of art with a reference to Hegel's question which "remains: Is art still an essential and necessary way in which that truth happens which is decisive for our historical existence, or is art no longer of this character?" (Heidegger 1993: 205).

Jameson finally answers this question in a way with his criticism of popart, which actually appears in direct opposition to what Heidegger and Hegel comprehended as art. Popart and the whole culture of postmodernism, as observed by Jameson, do not assure any unconcealment of the truth, they don't create any such representations. Pop art products are not "works" in the Heideggerian sense; they are simulacrums.

> Appropriately enough, the culture of the simulacrum comes to life in a society where exchange value has been generalized to the point at which the very memory of use value is effaced, a society of which Guy Debord has observed, in an extraordinary phrase, that in it 'the image has become the final form of commodity reification.' (Jameson 1991: 18)

Jameson is focused on the problem of the image and its representational ability, whereas, according to his observation, pop art fails to present the truth in the manner as Van Gogh's painting represents the peasant shoes by communicating the truth about the life, the everyday struggles and worries of the peasant. Warhol's paintings, according to Jameson, don't *stand for* any kind of truth or reality they would aim to unconceal. They are "empty" signifiers, bare images oriented only towards themselves.

The way Jameson understood Van Gogh's paintings is accordant with Guy Debord's comprehension of images in the society of the spectacle: "Everything that was directly lived has receded into a representation." (Debord 1989: 7). This does not mean that the images follow and depict reality. Debord placed stress on the fragmentation of views that takes place with the flood of images, which finally establishes a separate "pseudo-world". Therefore: "The spectacle is a concrete inversion of life, an autonomous movement of the nonliving." (Debord 1989: 7). The simulacrum is conceptualized as an illusion of reality, wherein the signifier no longer refers to reality. Yet, the criticism of images has relevant ground in Plato's criticism of images as substitutes for the real things. According to the Platonian perspective images are to be examined in relation to the truth. Images have a secondary status in this regard, as they are not real or true, therefore also not, as formulated by Susan L. Feagin "as powerful as the original, and so they are condemned as being weak imitations of the real thing, with a correlative tendency to confuse us about the nature of truth and reality" (Feagin 1995: 267). Yet Feagin contributed an important argument against this assertion that images are weaker than the original, namely she claimed that "substitutes extend the power of the original" by making the signified being present in different places at once, such as the emperor in every province, the Virgin Mary on every altar, ancestors in every vestibule, the president in every post office. Ludwig Feuerbach (1841) noticed the power of images, which is based on widening the gap between the signifier and the referent and which inverts the relevance of the original in favor of the copy: "But certainly for the present age, which prefers the sign to the thing signified, the copy to the original, fancy to reality, the appearance to the essence, this change, inasmuch as it does away with illusion, is an absolute annihilation, or at least a reckless profanation; for in these days *illusion* only is *sacred, truth profane*." (Feuerbach 1855: 10) This remark of Feuerbach that truth has become profane and the illusion sacred anticipated the notion of the simulacrum as conceptualized by Jean Baudrillard: "Neofiguration is an invocation of resemblance, but at the same time the flagrant proof of the disappearance of objects in their very representation: hyperreal. Therein objects shine in a sort of hyperresemblance /…/ that makes it so that fundamentally they no longer resemble anything, except the empty figure of resemblance, the empty form of representation" (Baudrillard 1994: 45).

As regards the showing of art works to the public, Jameson did not pay attention to the exhibiting environment in which art is to take place. Here, there is a relevant difference between the ever-present images and the spaces appropriate for showing art. These spaces create conditions for certain devotion and to perform a ritual to experience art, which holds some religious qualities of experiencing art. Walter Benjamin, one of the scholars of the Frankfurt school, linked traditional sorts of artwork, such as paintings, to the cult, magic and religion: "Originally, the embeddedness of an artwork in the context of tradition found expression in a cult. As we know, the earliest artworks originated in the service of rituals—first magical, then religious. And it is highly significant that the artwork's auratic mode of existence is never entirely severed from its ritual function" (Benjamin 2008: 24). Feagin considered paintings to be the ones to have the transformative power for the spaces to become ritualistic places: "paintings transform the space, and place, where they are, into one where certain ritual observances are appropriate" (Feagin 1995: 265). It could be added however that paintings also have a sort of transformative power for the observers, which is particularly effective if they are shown in spaces that establish better conditions to bring this power to effect. Precisely for this reason, to enable the enforcement of this power which is supported by the ritual, modernism established the gallery as a white cube. The white cube is still the appropriate space for showing Warhol's paintings. The ideology of the gallery as the white box was profoundly analyzed by Brian O'Doherty (first in 1976). The white, ideal space is an archetypical image of art in the 20th century. It comprises some of the sacredness of the church, formality of the court hall and mysticism of the experimental laboratory, and is set in an elegant design, which all together establish a unique chamber of aesthetics. The ideal gallery extracts any hint that testifies to its artfulness from the work of art. The work is in a white box isolated from everything that would draw the observer's attention away from the art on display (O'Doherty 2000). The purposeful gallery thus creates an ideological context that enables the objects exhibited there to become consecrated as the works of art. The art gallery also supports the ritual of experiencing art and of the contemplation of the works of art. In such a manner the artistic context establishes the conditions for the truth to happen through the work of art. It seems that such were also Jameson's expectations, as he was shocked by the flatness of Warhol's paintings.

Heidegger was aware of the proper way to present such works of art as he appreciated, which is to exhibit them. This means that how art is presented is not irrelevant. This setting up, however, does not mean a bare placing. Setting up a work is "an erecting that consecrates and praises", "the work opens up a *world* and keeps it abidingly in force" (Heidegger 1993: 169). Exhibition establishes conditions for the consecration of the works of art. Heidegger recognizes the analogy between

the sacred temples and art exhibitions. The spaces of art consolidate the sacredness of art and invite people to perform rituals of the appreciation of art. The work of art opens up a world at the exhibition as well as the temple opens up a world. "The temple-work, standing there, opens up a world and at the same time sets this world back again on earth" (Heidegger 1993: 168). The temple comprises sacredness: "By means of the temple, the god is present in the temple" (Heidegger 1993: 167).

One aspect of the end of art, addressed by Hegel and Heidegger, but also by Jameson, can be related to the withdrawal of the sacredness related to art and its enjoyment. Jameson actually determined this has happened with pop art. Consumerism as a theme entered the world of art. In addition, the works of art have themselves become fetishized commodities par excellence (Tratnik 2019). This is another effect of consumerism that Jameson did not recognize. But art also moved out of the context of art into other public spheres and into the centers of consumption. In order that this could become possible, art had to change its ontological status from being the representational means for unconcealing the truth to becoming an action, a performance, an intervention. Not only have changes occurred to the institution of art, the religious moment has found its place in the increasingly growing centers of consumption. These facts open up space for quite different considerations about consumerism and art in late capitalism.

2. *Art as Intervention into the Temples of Consumption*

In his study of postmodernism as the cultural logic of late capitalism, Jameson noted that it was from architectural debates that his own conception of postmodernism initially began to emerge (Jameson 1991: 2). In architecture, postmodernism stages itself as aesthetic populism, which has at least the merit to efface the high-modernist frontier between high culture and commercial culture (Jameson 1991: 2). Although interested in consumerism as the central feature of late capitalism, Jameson, however, does not examine how the means of consumption seduce the consumer and how shopping malls organize people's behavior. He is not interested in the consumer's experience of shopping, the milieu in which consumption takes place, consumers' psychology, their shopping addiction, and the ideology at work in the places of consumption. All these are the actual grounds that establish the cultural logic and the success of late capitalism. At about the same time that Jameson first published his theory of postmodernism (in 1984), American theologian Ira G. Zepp analyzed shopping malls as ceremonial centers (Zepp 1986). Zepp acknowledged that in the urban USA (of that time, in the 1980s) people continue to seek community, construct centered spaces and ritualize their lives, this time through shopping malls. These comprise mythic geometry, architectural rhetoric and offer a meaningful variety of human activities that take place there. Even earlier than Zepp, in 1980, cultural anthropologist Alexander Moore analyzed Walt

Disney World as a ritual space and a playful pilgrimage center. Moore ascertained that Walt Disney World organizes behavior by combining play and ritual, which "comprise a metaprocess of expressive behavior rooted in our mammalian past." (Moore 1980: 207) The ritualization of the lives of the majority of people in late capitalism has not only fundamentally evolved around consumerism. The means of consumption have become the "cathedrals of consumption", ascertained George Ritzer and he points out the "quasi-religious, 'enchanted' nature of such new settings" (Ritzer 2005: x). He establishes that people make "pilgrimages" to these places, "in order to practice our consumer religion" (Ritzer 2005: x). These means of consumption are structured to have enchanted, even sacred, religious character—sociologist Peter Corrigan acknowledged that "the Church and industry can draw upon the same awe-inspiring techniques" and recognized "department stores as similar to cathedrals" (Corrigan 1997: 56). They are immense, vast gigantic, with huge galleries and staircases inside the buildings and enable to look down into the vast and bustling throng. Corrigan compares the effects of the architecture of the sites of consumption with architecture of the church: "size is a characteristic of many buildings that are designed to awe small human creatures" (Corrigan 1997: 55). The other relevant idea emphasized by Corrigan regarding shopping malls is that they provide everything. He links this idea to the "broader notions of power: rulers who really can provide everything will forever have people in their debt, and a department store may well borrow some of the same effect" (Corrigan 1997: 55). Suddenly such a place "almost magically abolishes all thought of deficits or shortages" (Corrigan 1997: 55).

Ritzer further establishes that shopping malls have become "more than commercial and financial enterprises; they have much in common with the religious centers of traditional civilizations" (Ritzer 2005: 8). People are enthusiastically part of consumer society and these settings offer the "greatest spectacle" (Ritzer 2005: xi). Because we can easily grow bored, the consumptions settings compete to see "which one can put on the greatest show" (Ritzer 2005: xi). Not only shopping malls, but different types of settings "are rushing to emulate the cathedrals of consumption", such as universities, fast food restaurants, souvenir shops, video arcades (Ritzer 2005: x), and even megachurches which offer aerobic classes, bowling alleys, counseling centers, and multimedia bible classes (Ritzer 2005: 23).

Zygmunt Bauman analyzed shopping as a rite of exorcism. People are "running after pleasurable—tactile, visual or olfactory—sensations, or after the delights of the palate, promised by colourful and glittering objects displayed on the supermarket shelves" or "sensations promised by a session with a counselling expert. But they are also trying to find an escape from the agony called insecurity" (Bauman 2000: 81). The crowds gather in the temples of consumption, but not in order

to talk and sociate, points out Bauman. Encounters in a crowded space are brief and shallow. People don't establish deeper or more complex relationships. However, crowded the place may be, it is not collective, reflects Bauman. "To deploy Althusser's memorable phrase, whoever enters such spaces is 'interpellated' qua individual, called to suspend or tear up the bonds and shed loyalties or put them on a side burner" (Bauman 2000: 97). This fundamental link between the means of consumption and ideology is crucial to understanding the processes of subjectification and desubjectification taking place via the means of consumption, which is itself an apparatus. As demonstrated by Michel Foucault and Giorgio Agamben, apparatuses in a disciplinary society seize the bodies in the very process of their desubjectification. In order to produce new subjects, the apparatuses first split the subject through the negation and at the same time the assumption of the old. The apparatus of the prison, for instance, produces "the constitution of a subject and of a milieu of delinquents, who then become the subject of new—and, this time, perfectly calculated—techniques of governance" (Agamben 2009: 20). In a quite similar manner the means of consumption produce the constitution of a subject and a milieu of consumers who become the subject of the calculated technique of governance, altogether with the ultimate objective to accumulate capital and to increase the economic and political power of the social elite.

In 2003 Laibach, the founding art and musical group of the art collective Neue Slowenische Kunst, performed a visit to the shopping mall City Park in Ljubljana. As part of their outfits the group wore a kind of military uniform particularly reminiscent of the Wermacht uniform. The group became known for provoking the public during the disintegration of the Eastern bloc with ambiguous performances with which they addressed the functioning of totalitarian systems. Their interventions were a torment for a society with collective memory of the German and Italian occupation. Within the performance *Einkauf* (*Shopping*) they simply took a shopping cart and walked in a military style around the mall, doing their shopping in their uniforms with serious expressions on their faces. Their performance was uncomfortable for the people present in the given context. The consumers felt threatened and security guards were unsure whether or not to stop the intervention. The invisible envelope of the Arcadian environment designed for enjoyment and relaxation was suddenly broken. The action *unconcealed* the fact that the consumerist space is clearly politically structured in a totalitarian manner, as it does not tolerate any penetration of other ideas, behaviors, rites, or ideologies. The performance showed the parallels between consumer ideology, as well as the functioning of its means of consumption, and political totalitarian systems. The seemingly non-political consumerism is proved to be fundamentally political. Art has, in this case, shocked the present public who was not expecting to come in contact with art in that context. Its function was

this time to awaken consumers and to unconceal the political structure of the temple of consumption.

Temples of consumption offer "the comforting impression of belonging" (Bauman 2000: 99). Bauman pictured the places as a "floating boat", a "self-contained 'place without a place'", which is a "*purified* space". The place had been cleansed of variety and difference; the differences inside "are tamed, sanitized, guaranteed to come free of dangerous ingredients—and so be unthreatening", so that "what is left is pure, unalloyed and uncontaminated amusement" (Bauman 2000: 99). The isolative and the excluding character of the temples of consumption is complete:

> The place is well protected against those likely to break this rule—all sorts of intruders, meddlers, spoilsports and other busybodies who would interfere with the consumer's or shopper's splendid isolation. The well supervised, properly surveilled and guarded temple of consumption is an island of order, free from beggars, loiterers, stalkers and prowlers—or at least expected and assumed to be so. (Bauman 2000: 98)

Consumption is related to destruction of critical potential and moral indifference. In 2006 the artist Sašo Sedlaček hit in the core of this problem with his project *Beggar*. The *Beggar* is a robot made at home from recycled material, designed to help socially marginalized people. Sedlaček noticed that the huge places of consumption are exclusionary, as there are no homeless people to be found there, whereas poverty is an increasing social problem. Despite replacing historic city centers, these places have not fully assumed the function of an open public space for all (Tratnik 2009: 18). The *Beggar* was let into the City Center in Ljubljana as a robot collecting money for the homeless. It collected much more than homeless people collect in the same time frame on the street, selling their newspaper *Street Kings* which is meant to be an alternative to direct begging. This says a lot about the compassion people have toward the poor compared to the sympathies we have towards digital technology. If on the one hand people feel uncomfortable when confronted by poverty and homelessness which make invoke feelings of fear, on the other hand they are attracted to the mechanical or even more with the digital gadgets that they can play with.

The means of consumption rely on the mechanism of seduction. Bauman recognized that the heavy, Fordist-style capitalism, passed over to the light, consumer-friendly capitalism. If the first was the world of the rulers, law-givers, and supervisors who directed other people, the latter preserved authorities but now authorities coexist. The authorities of light capitalism no longer command, but they ingratiate themselves with the chooser; they tempt and seduce (Bauman 2000: 63). Yet in a consumer society everything is a matter of choice except the compulsion to choose. This compulsion grows to an addiction and is thus no longer perceived as a compulsion (Bauman 2000: 73).

After the fall of the Berlin Wall, the former communist European countries started to join the European Union one after another, as to an

alliance of the European countries with a collective political governance and a foundation in capitalism. In the former communist countries, the shopping centers were quickly built during the transition to capitalism. The phenomenon of consumerism at once struck the population in these countries. People who experienced a shortage of goods under the previous regime and used to smuggle them from other countries where they would go on shopping trips, were suddenly overwhelmed by the richness of the offer, the surplus of commodities that suddenly appeared near their homes. They were enchanted by the means of consumption and the complete experience that shopping at once became. The population, disappointed with the previous ideology, was subject to uncritically accept the arrival of the new capitalist ideology. The directors Vít Klusák and Filip Remunda showed how high the level of seduction and even addiction to shopping was for the Czech population in 2004, with their film project *The Czech Dream*. They studied the components required for the successful establishment of a shopping mall, apart from its construction, that would draw the consumers to the defined location. Klusák and Remunda built the whole promotion of the coming center; they made out their own appearance, conducted the advertising campaign and even produced a theme song to emotionally attract the consumers. Finally, people arrived many hours before the expected opening of the mall, but the location was just a rural field and instead of wandering in the shopping mall they got a chance to take a walk in the countryside. Many of the people who had arrived expecting to have a shopping experience felt angry. Afterwards, the project triggered rich public discussions as rarely seen regarding consumerism, the role of art and public financing of art, as well as on the political question of joining the European Union, which was current at the time.

In 2003 Sašo Sedlaček collected the advertising leaflets that he had been receiving from the shopping centers, and invited the public to build bricks out of this material in the gallery space (Kapelica Gallery). With a group of colleagues, he then conducted an action: with those bricks they built a wall with which they closed off the entrance of the City Park in Ljubljana. He accompanied the intervention, which he called *Just Do It!*, by stating that he was only giving back what he had received and had not asked for.

These cases of performative art unconceal the truth of capitalism, its refined hidden mechanisms and ideology, as well as its (side) effects, such as the increase in poverty within the population, ecological pollution due to the hyper-production of goods, shopping centers, and promotional material. Art is, in this case, not experienced through the mechanism of contemplation. Its function is to stimulate critical thinking. This is possible with an intervention of art into the marrow of capitalism, into its means of consumption which is the capitalist means of enchantment. This is an intervention into the Church of capitalism, the sacred environment, where the consecrated ritual of seduction takes place.

Conclusion

Jameson's mourning for contemporary art that does not fulfil expectations to be critical towards capitalism is grounded in his ontology of art. His comprehension of true art is built upon Martin Heidegger's concept of art. For Heidegger, art is expected to communicate truth. Heidegger's ontology of the work of art rests upon Hegel's ontology of art, according to which art has the status of representation in the sense that it is established as a secondary reality which is set in relation to the first. As such it functions as a medium, intervened by the artist, which can facilitate the unconcealment of the truth. Jameson's ontology of art did not pay regard to the rising phenomenon of the temples of consumption which are the crucial point for the examination of culture in late capitalism. Jameson did not fully consider the passage of capitalism from the end of the 19th century or the beginning of the 20th century to contemporaneity, or from what Zygmunt Bauman called the "heavy", Fordist-style capitalism, to the light, consumer-friendly capitalism which is signified by the growth of temples of consumption. Jameson, furthermore, did not consider the great changes that art had gone through in the second half of the 20th century, from the representational to the performative mode of art. If he considered art as performative and had searched for its task in relation to the interpellation of the consumer, supported with the specially-designed environments as centers of experience, he might have imagined the possibility of critical art in the context of late capitalism much differently.

References

Agamben, G. 2009. *What is an Apparatuss?* Stanford: Stanford University Press.

Baudrillard, J. 1994. *Simulacra and Simulation*. Michigan: The University of Michigan Press.

Bauman, Z. 2000. *Liquid Modernity*. Malden: Polity.

Benjamin, W. 2008. *The Work of Art in the Age of Its Technological Reproducibility, and Other Writings on Media*. Cambridge and London: Harvard University Press.

Corrigan, P. 1997. *The Society of Consumption*. London: Sage.

Debord, G. 1989. *Society of the Spectacle*. London: Aldgate Press.

Feagin, S. L. 1995. "Paintings and their Places." *Australasian Journal of Philosophy* 73 (2): 260–268.

Feuerbach, L. 1855. *The Essence of Christianity*. New York: Calvin Blanchard.

Hegel, G. W. F. 1975. *Hegel's Aesthetics. Lectures on Fine Art*. Oxford: Oxford University Press.

Heidegger, M. 1993. "The Origin of the Work of Art." In *Basic Writings*. San Francisco: Harper: 139–212.

Jameson, F. 1991. "The Cultural Logic of Late Capitalism." In *Postmodernism, or, the Cultural Logic of Late Capitalism*. London, New York: Verso: 1–54.

Mandel, E. 1972. *Der Spätkapitalismus*. Frankfurt am Main: Suhrkamp.

Moore, A. 1980. "Walt Disney World: Bounded Ritual Space and the Playful Pilgrame Center." *Anthropological Quarterly* 53 (4): 207–218.

O'Doherty, B. 2000. *Inside the White Cube: The Ideology of the Gallery Space*. Los Angeles: University of California Press.

Ritzer, G. 2005. *Enchanting a Disenchanted World, Pine Forge Press. Revolutionizing the Means of Consumption*. 2nd edition. Thousand Oaks: Pine Forge Press (Sage).

Tratnik, P. 2009. "Brkljanje, ki sproti rešuje probleme. Pogovor s Sašem Sedlačkom." *Maska* XXIV (119–120): 14–19.

Tratnik, P. 2019. "Umetnost kot investicija: institucija umetnosti v službi umetnostnega trga." *Annales: anali za istrske in mediteranske študije, Series historia et sociologia* 29 (3): 473–484.

Zepp, I. G. 1986. *The New Religious Image of Urban America. Shopping Mall as Ceremonial Center*. Westminster, MD: Christian Classics, Inc.

Croatian Journal of Philosophy
Vol. XXI, No. 61, 2021
https://doi.org/10.52685/cjp.21.1.4
Received: July 29, 2020
Accepted: November 23, 2020

Fashion as an Aesthetic Form of Life: A Wittgensteinian Interpretation

ELENA ABATE

Fashion is an aesthetic practice which concerns the ordinary sphere of our life: it is associated with everydayness and it is a source of endless aesthetic experiences. The purpose of this paper is to validate a new perspective on fashion based on Wittgenstein's later aesthetic conception. In Philosophical Perspectives on Fashion (2017), Matteucci introduces the idea of combining the Wittgensteinian concept of "form of life" with fashion. In accordance with this thesis, the paper aims at showing how fashion is constituted as a "form of life". Specifically, I shall argue that fashion is an "aesthetics form of life" which structurally employs a language of an aesthetic type —one with a specific grammar (or set of rules) of its own. I claim that there is in fashion a contact point between the grammar of language and socially encoded aesthetic responses: fashion follows slavishly its own grammar, through its cyclical seasonality, while at the same time tending to creatively reinvent itself. Thus, anyone who daily commits to the practices of fashion acquires sensitivity to its rules, contributing to a social dialectic of identification/diversification typically belonging to fashion itself. Finally, on the basis of the claim that fashion is a "form of life", and indeed since fashion is primarily an aesthetic practice, I claim that Wittgenstein's aesthetic notions can coherently be related to fashion as well: concepts such as 'aesthetic reaction', 'gesture', and 'correctness' will be shown to be crucial to an analysis of the aesthetic phenomenon of fashion.

Keywords: Fashion; dressing; Wittgenstein; form of life; grammar.

1. *Introduction*

Within the discipline of aesthetics, the appearance of fashion as a field of enquiry in relatively recent. The abundancy of studies related to fashion in literature, sociology, psychology, economy and anthropol-

ogy is not nearly matched by the relatively exiguous number of studies devoted to the topic in philosophical aesthetics[1]. Nevertheless, as Ian King states, fashion can afford us a new opportunity to understand contemporary aesthetics, provided that our understanding of it will be detached from "the discussion in the aesthetics literature [that has been] normally theoretical, normative and reliant predominantly on understandings taken from fine art" (King 2017: 2). A different understanding or approach to fashion should take into consideration its practical and ordinary dimension, focusing on the daily activities and everyday objects that fashion itself involves. Before trying to spell out what I take to be the first essential step toward a different understanding of fashion, I wish however to define what I mean by "fashion."

The term "fashion" is often associated with an institutionalized system made up by groups, organizations, producers, events and practices, other than being simply associated with dress or clothing (cf. Kawamura 2005: 43). However, it seems rather arbitrary to exclude *a priori* from a definition of "fashion" some of the aspects of the complex and multifarious phenomenon we call "fashion", or at any rate to marginalize them. At the same time, in the condensed space of a paper, it is impossible to deal with most of them. Without claiming to be exhaustive, I shall thus restrict the field of application of my definition of "fashion" to two aspect aspects of the phenomenon of fashion, which are meant to support my arguments in what follows. First, I take "fashion" to signify an everyday practice with which we confront daily. Specifically, I mean the ensemble of actions involved in the practice of dressing ourselves up.[2] Secondly, I take "fashion" to signify an ensemble of objects and activities permeated by a deep aesthetic dimension (which I intend to articulate in the present paper), insofar as fashion involves embodied collective experiences through which is possible to create—and move into—a meaningful aesthetic space. Given my definition, I thus take fashion to be both tangible and intangible, because it implies experiential practices—such as dressing ourselves, buying and wearing items, etc.—while at the same time concerning an immaterial or non-strictly-material domain, made up of aesthetic proprieties, relations, reactions, expressions, and values.

2. *Fashion as ordinary aesthetic practice*

2.1 *Fashion within aesthetics*

The interrelation between fashion and aesthetics can be traced back to the 19th and 20th centuries, when philosophers and literati[3] displayed

[1] While the sociological, psychological, economic, moral, political, anthropological and more other approaches to fashion are of great importance, in this paper, for reasons of space, I will concentrate on the aesthetic dimension of fashion alone.

[2] This sense of fashion might be considered closer to the concept of 'clothing' or 'dress', which are connected to fashion in a wide sense in turn.

[3] Cf. Matteucci 2019. Matteucci quotes philosophers (such as Simmel, Spencer, Benjamin and Fink) and men of literature (namely Balzac, Baudelaire, d'Annunzio, Carlyle).

an interest in the aesthetic implications of fashion for the first time. Yet, nowadays the aesthetic dimension of fashion seems to be obvious; so obvious, I believe, that it often escapes consideration. In this section, my aim is to describe the first theoretical steps that have been taken toward a rediscovery of how fashion relates to reality creating deeply aestheticized fields that permeate our ordinary life.

To give a clear account of why fashion has to be considered an aesthetic phenomenon, I shall first of all mention Iannilli's (2017) reasons concerning why fashion should be considered as one of Everyday Aesthetics' key topics. The first reason—described as "empirical"—concerns the production of the fashion items which affect our everyday life: fashion provides an ensemble of objects—e.g. clothing items and accessories—to which we refer through descriptions that pick out aesthetic qualities and properties[4]. The second reason is described as "theoretical": fashion shapes our ordinary experience and "emphasizes specific structures of the latter that cannot be neglected from a purely aesthetic point of view, since they indeed result in a very particular configuration of the aesthetic in everyday life" (Iannilli 2017: 231). Further, fashion contributes to the processes of aestheticization by creating immersive experiences of continuous consumption. Indeed, it is undeniable that fashion presents an aesthetic dimension, especially within our 'post-industrial' society, in which aesthetics "can no longer be confined within an ideal and isolated sphere such as the system of fine arts governed by the industrial society" (Matteucci 2016: 10).

Notwithstanding all of this, for long time, everyday practices and objects—such as fashion—have been considered too ephemeral or superficial to have an impact on aesthetic reflection. Only comparatively recently philosophers have turned their attention to our deeply aestheticized ordinary contexts (to which are dressing practices obviously belong). In fact, as Dewey[5] claimed, the aim "is to restore continuity between the refined and intensified forms of experience that are works of art and the everyday events, doings, and sufferings that are universally recognized to constitute experience" (Dewey 1934: 3). Everyday aesthetics (EA) aim to respond to an increasingly severe restriction of the aesthetic domain over time, which started with Kant's aesthetic account; focusing the aesthetic inquiry almost entirely on the definition of beauty and the sublime, as Kant did, caused in fact a significant art-centred shift in the discourse of aesthetics qua discipline.

[4] For example, one can define a certain dress as 'beautiful' or 'ugly' or 'cute', etc.

[5] According to Yuriko Saito's entry in the Stanford Encyclopedia of Philosophy, Dewey is considered the forerunner of Everyday Aesthetics: focusing of the aesthetic experience of an artwork (*Art as Experience,* 1934), he encouraged to concentrate on ordinary experience, rather that artistic products, insofar the starting point of aesthetics is the aesthetic "in the raw" (Saito 2019).

Clearly, however, our everyday life deserves aesthetic scrutiny as well. And so, as Ratiu clarifies, "the scope of aesthetics [is today] expanded to include areas of everyday life previously neglected" (Ratiu 2013: 7). Specifically, daily objects and quotidian activities, such as dressing up, can now claim to belong legitimately to the scope of aesthetics, insofar as they "they contribute constantly and crucially, in everyday life, to the configuration of the taste of individuals and to the understanding of the environments that surround us and with which we interact." (Matteucci 2017: 111–112). Indeed, "the recognition of fashion's dignity as a philosophical and specifically aesthetic research topic seems to imply the necessity for a broadening of the domain of aesthetics itself, for it traditionally follows an art-centred and/or nature-centred paradigm of investigation" (Iannilli 2017: 230).

Already considering the short summary above, I think it is fair to say that the question about the correlation between fashion and the realm of aesthetic is often misinterpreted. Instead of asking "why does fashion have an aesthetic dimension?", it is more appropriate to question why the discipline of aesthetics did not consider fashion as an important topic to discuss in the first place. This sort of considerations makes clear that a reconsideration of fashion within an aesthetic framework is urgently required.

2.2 *A new aesthetic interpretation of fashion*

Concerning the aesthetic dimension of fashion, it is crucial first of all to define in which way fashion intertwines with the ordinary element. In order to give such description of fashion, I wish to introduce Matteucci's notion of hypo-aesthetic level[6], which he uses to illustrate the "presence of aesthetic elements in everyday life" (Matteucci 2016: 14).

The hypo-aesthetic level concerns the diffusion of aesthetic elements within everyday life—those elements, that is, that constantly shape our interaction with reality in an aesthetic way, working underneath the surface. As a matter of fact, human beings have always shaped their reality through aesthetic devices, for example words and images used in a non-denotative sense. The deep historical and psychological root of such aesthetic phenomena could also instruct us about the recent structural changes in our ways of experiencing reality, in which fashion plays a non-marginal role. This genetic analysis, in turn, would pave the way for the thesis that fashion constitutes—at least nowadays—a peculiar "form of life", the elements of which were already dormant in previous (non-thoroughly-aestheticized) forms of life.

From the perspective of the hypo-aesthetic level, fashion is to be

[6] See Matteucci 2016. In contrast to the hypo-aesthetic level, Matteucci also introduces the notion of 'hyper-aesthetic level', which refers to those aesthetic elements in everyday life that can trigger experiences with an abundance of aesthetic content. Although relevant connections could be made between this level and fashion, I will not have enough space in this paper to investigate this correlation.

understood as a deep routine that "remains with people over time" (Buckley and Clark 2012: 19). In our daily life, we use and remodel fashionable clothes, again and again. Nevertheless, we usually refer to fashion in terms of regular stylistic innovation supported by an institutionalized system of production. Through this standard notion of fashion (qua innovation, perpetual change, etc.), we perceive the system of fashion as extraordinary, extravagant, uncommon: "in fact according with Buckley and Clark (2012: 20) while the extraordinariness of high fashion has been clearly visible, ordinary fashion has been resolutely invisible." However, fashion can be both: ordinary, in its everyday dimension, and extraordinary, as fashion itself sets the seasonality and the novelty. People constantly dress themselves and, in this way, depict their interpretation of the cycles of fashion. We experience fashion every day without noticing that we are experiencing it: in fact, as a proof of this inner behaviour, we dress ourselves up every day.

"The problem, as Sheringham points out (2006: 22) [is that] the everyday is beneath our attention. It is what we overlook." Dressing ourselves is such a simple task that in every moment of our daily routine it might be perceived to be obvious, precisely because the ordinary escapes notice.[7] In fashion the perception of everydayness is hard to locate because fashion is mostly identified with modernity, fastness, fleetingness. However, if we pay attention to both aspects of fashion outlined above, we can easily see that fashion is indeed able to create a conjunction between modernity (intended as velocity and variableness) and everydayness. Thus, fashion offers the possibility to set practices regarding our everyday life, combining creativity and repetitiveness: it provides the possibility of reinventing one's own image repeatedly within everyday life.

Considering what we have said so far, fashion appears as an ordinary aesthetic practice that concerns also a particular kind of objects that can acquire aesthetic properties in everyday contexts. Nevertheless, fashion is still a controversial topic within the domain of the Everyday Aesthetics, and indeed as a general aesthetic phenomenon. The problem is that fashion is and has been treated poorly with respect to other human phenomena we study, as it is considered lacking in a solid theoretical basis. The theoretical unreliability that is historically associated with fashion is due to the fact that, at first glance, it presents itself as "a bundle of problems that join together in an irregular manner" (Matteucci, 2017: 13).

There are, however, encouraging signals of change with respect to this problem. Despite the lack of systematic philosophical enquiries concerning fashion, Matteucci has recently provided a critical examination of fashion (cf. Matteucci and Marino 2017) comparing four dif-

[7] As Wittgenstein writes: "The aspects of things that are most important for us are hidden because of their simplicity and familiarity. (One is unable to notice something—because it is always before one's eyes.) […] And this means: we fail to be struck by what, once seen, is most striking and most powerful" (PI, § 129).

ferent aesthetic paradigms to fashion, in order to understand which one can describe fashion in the most accurate manner. Here, I wish to show that one of these paradigms, namely the one that reads the phenomenon of fashion in the light of Wittgenstein's concept of "form of life", can be the most fruitful one, if our aim is that of giving an aesthetic characterization of fashion. Thus, the purpose of my arguments below will be to demonstrate that the concept of "form of life" applied to fashion might constitute a viable paradigm for fashion, if inflected in an aesthetic manner, in terms of what I shall call "aesthetic form of life." Specifically, I will argue below that this paradigm suits fashion as an everyday practice, intended as the repetitive act of dressing ourselves up in daily life, in certain manners, which depend on the specific occasions.[8]

3. *Form of life and rules*

Matteucci's comparison uses the analytical aesthetic theory by Wollheim in *Art and its Objects* (1968): outlining what is to be understood as art and as an aesthetic, Wollheim compares art to a "form of life" using this concept in the meaning with which Wittgenstein himself explained his conception of language and language games. By "form of life" we mean therefore the set of habits, intrinsic experiences, indeed a language and its uses. Equally, to describe the aesthetic field as a "form of life" implies considering perceptive and cultural experiential practices in which the subjects involved express themselves by drawing a horizon of shared taste. However, expressing one's own horizon of taste does not mean establishing a static set of signs or indexes to represent things, that is, following a semiotic modus operandi. On the contrary, drawing a horizon of taste is to bring out the physiognomy of things in common ways, one expressing familiarity. A system of familiar aesthetic relations, in this sense, is manifested through the various forms of taste, such as art and even fashion.

I argue that the best way to describe fashion's mechanism is comparing it to the late Wittgenstenian concept of language, and in particular with the "form of life" concept. By 'form of life' we mean therefore the set of habits, intrinsic experiences, indeed a language and its uses. Wittgenstein describes this concept as an *Übereinstimmung* (agreement, concordance) in the sense of producing consensus and regularity. The intersubjective agreement within a form of life is situated in language interpreted as a universal medium and as a place of consensus and possible constitution of experience in the world (Borutti 1993: 1). Language as a form of life is a condition of a possible community.

The concept of "form of life" is tied to the concept of language, as Wittgenstein points out in *Philosophical Investigations* in § 19: "to

[8] These occasions include, for example, going to work, to an interview, to a wedding, to a theme party, to a dinner or a business lunch, to the park, and so on.

imagine a language means to imagine a form of life" and "speaking a language is part of an activity, or a form of life" (PI, § 23). In the perspective of *Philosophical Investigations*, language is made of several language games at the same time. Language games are defined by Wittgenstein as "objects of comparison which are meant to throw light on the facts of our language by way not only of similarities, but also of dissimilarities (PI, § 130)" and are used by Wittgenstein with the intention of shedding clarity onto language. Language games are models that expand our way of looking at language and allow us to observe its multiplicity. The rules of a game are not strict: "the rule may be an aid in teaching the game (PI, § 54)": you can learn a rule either by observing a game or by playing it. In the latter case, a player can understand the rules of a game directly through practice.

In the case of language, this means that one can understand the meaning of different words in a language game and the specific rules governing their use, as the game allows access to a field of application of the words themselves (the use of the words in language). In fact, Wittgenstein affirms also that "without these rules, the word has no meaning, and if the rules change also the meaning changes" (PI, § 552). By following the rule, one can understand, at the same time, what the rule is and how to apply it.

Furthermore, "also 'obeying a rule' is a practice. And to *think* one is obeying a rule is not to obey a rule. Hence it is not possible to obey a rule 'privately': otherwise thinking one was obeying a rule would be the same thing as obeying it" (PI, § 202). In fact, to follow a rule is a public practice as it implies the recognition of the rule by people who follow the same rule. It is impossible to follow a rule privately because following a rule is itself a practice, which requires approvals, disapprovals, gestures, orders that strengthen the rule, and so on. According to Wittgenstein, these are "grammatical annotations" on the expression of following a rule that concerns habits upon which humans agree.

4. *Wittgenstein's perspective on aesthetics*

Within an aesthetic perspective, Wittgenstein dealt with aesthetics by dealing with problems concerning the meaning of aesthetic words. Before Wittgenstein, the concept of beauty was the starting point of aesthetic reflection. However, Wittgenstein noticed that all sorts of confusions and misunderstandings arise from the analysis of the form of traditional aesthetic utterances, i.e. the Kantian type of proposition "this X is beautiful" (Johannessen 1996: 24–25). In fact, we are misled into thinking that it is possible to infer structural features of the world thanks to the structural features of the proposition: "according to Johannessen (1996: 25) we are in fact tempted to reason from language to reality." This temptation leads us to ignore all the other kinds of aesthetic judgments that, according to Wittgenstein, concern aesthetic features far more important than beauty, and that which we call beautiful.

Indeed, Wittgenstein breaches into the tradition of aesthetic enquiry claiming that it is not necessary to find a correct and universal definition of beauty, because, as with the concept of art, the borders and the application of the concept of beauty are essentially vague.

Furthermore, whenever Wittgenstein refers to aesthetic concepts, he accepts their radical indeterminacy: he claims that it is not worthful to ask for the precise definition and boundaries for aesthetics concepts because their very nature is vague.

In *Lectures and Conversations on Aesthetics, Psychology, and Religious Belief* (LC, 1966) he mentions his disinterest in 'Aesthetic as a science', i.e. the science of the beautiful. Thereby, "Wittgenstein (LC, 2:1) starts by investigating what could be meant by 'Aesthetics'", claiming that the aesthetic field "is very big and entirely misunderstood" (LC 1:1). In fact, he points out that it is useless to focus only on terms like "beautiful" or "ugly" because in aesthetics there are a great amount of interjections and reactions to artworks or natural beauty.

Primarily, Wittgenstein concentrated on the use of aesthetic expressions and their linguistic form. He wanted to explore how and where aesthetic judgments are employed in daily life. He counters the traditional aesthetic discourse paying attention to what happens in real life, "claiming that what we do is to bring words back from their metaphysical to their everyday use" (PI, § 48). He started by reflecting on aesthetic phenomena; specifically, he was interested in situations where aesthetic disagreement arises about a given aesthetic matter. In a way, aesthetic enquiry evolves into a discussion concerning a disagreement over some artistic expression. The matter is localizing the source of what Wittgenstein called "aesthetic puzzlement": when we encounter certain artworks, we experience disquiet or aesthetic discomfort and, at the very same time, we feel confused on the source of our experience (Johannessen 2004: 17).

According to Wittgenstein, the concept of aesthetics points to whether something is working or not, if it is pleasant or unpleasant, beautiful or ugly: in other words, aesthetics means to perceive that something has the right expression (or not), the right gesture (or not). Aesthetics is about understanding why something is right or wrong, and indeed about providing reasons for this. Only by understanding why something is right (or not) it is possible to change the way a person perceives external things. Accordingly, research in aesthetics has firstly to deal with "aesthetic uncertainty", indeed with situations in which the subject feels dissatisfied or disgusted by something, but he does not know why he has this kind of reaction.

From the discussion above, however, it will also be clear that aesthetics has also to deal with an ensemble of aesthetic reactions involving expressions and gestures aimed at the object that prompts the reaction. As it was for language games, aesthetic reactions and aesthetic judgments display family resemblances: expressions like "beautiful" or "awful" are used in our linguistic practices as interjections, particularly

when we experience an artwork. In fact, as Wittgenstein points out, aesthetic reactions are of great significance in addressing the concept of aesthetics (cf. LC, 2:10); and from aesthetic reactions, such as interjections, will derive aesthetic language games (cf. LC, 1:5). Through aesthetic reactions, we can go back to the reason—as opposed to the cause—which motivates our aesthetic reactions. "According to the 1933's lessons (MWL, 9:30) a reason in aesthetics is a reason for having this word in this place rather than that; this musical phrase rather than that. Reason = justification."

The aim is trying to resolve our aesthetic puzzlement by giving aesthetic explanations. Indeed, as Wittgenstein (MWL, 9:27) pointed out: "The question of Aesthetics is not: Do you like it? But, if you do, why do you?."

The kinds of reasons we appeal to when we try to resolve our aesthetic perplexity are not causal explanations. Aesthetic impressions and reactions cannot be explained by external-causal matters: "There is a 'Why?' to aesthetic discomfort not a 'cause' to it" (LC, 2:19). Wittgenstein rejected considering aesthetics a branch of psychology because it is not possible to explain aesthetic experience through causal explanations (Glock 1996: 33). "He firmly pointed out in his 1933's lessons (MWL, 9:32) that the reasons have nothing to do with psychology." Indeed, Wittgenstein's aim is to formulate a grammatical explanation.

We can explain our aesthetic reactions through a better understanding of the work of art itself, which proceeds linguistically. Aesthetic explanation is descriptive; therefore, it is possible to distinguish between our aesthetic reactions and their reasons: "What reasons could I give for being satisfied? They are in the nature of further descriptions" (MWL, 9:31).

Wittgenstein talked about rules also within the aesthetic frame. In fact, the use of language is thus the object of study of aesthetics. As we have seen, in his 1933 lectures (MWL, 4b), Wittgenstein already pointed out that beauty has no central relevance in aesthetics. In fact, it is rare to use terms such as "beautiful", "magnificent", "fine", "ugly" in aesthetics. Words or expressions such as "beautiful" are not so important in our discussion of artworks: "beautiful" is only one type of word, that we can encounter at best sometimes. Later, during the 1938 lessons, Wittgenstein reiterated his account, highlighting that only he who has a lack in aesthetic competence uses aesthetic adjectives such as "beautiful" or "magnificent".

Indeed, our aesthetic competence is tied to aesthetic rules: "Wittgenstein (LC, 1:11) asked in fact what rule are we using or referring to when we say: 'This is the correct way'? If a music teacher says a piece should be played this way and plays it, what is he appealing to?." This question is not as easy as one might think, and it requires an account of aesthetic rules in the first place, and their correlation with aesthetic judgements (such as the teacher's one).

Familiarity with a set of aesthetic rules is a necessary condition for giving a competent aesthetic judgment according to the criteria of aesthetic correctness or aesthetic incorrectness (with respect to a given rule). It is also fundamental in aesthetic reasoning, and thus needed to make someone see what you see in art. Aesthetic judgment works in the same way; in fact, for Wittgenstein, what gives meaning to our aesthetic judgment are not merely the words contained in it, but the complex of cultural contexts and human activities in which we can actually find the usage of these words (Shusterman 1986: 98).

The essential point is this: the more our knowledge of aesthetic rules is accurate, the more our aesthetic judgment will be appropriate. In fact, as Wittgenstein claimed, aesthetic rules are vital to aesthetic judgments: he clarified that "if I hadn't learnt the rules, I wouldn't be able to make the aesthetic judgement. In learning the rules, you get a more and more refined judgement. Learning the rules actually changes your judgement" (LC, 1:15). This means that, although aesthetic judgment is constrained by rules at any given time, it is also possible to acquire a "feeling for the rules" (Novitz 2004: 58), which opens the possibility to refine our aesthetic judgements over time. Moreover, aesthetic rules are embodied in our culture, which of course does change over time as well.

The meaning of an aesthetic judgment, and indeed of all the aesthetic activities that surround it, can be found within the socio-cultural context in which the judgement is used, and thus, ultimately, in its role in our "way of living". In a way, our paradigms of aesthetic evaluation are as obscure or complicated as is their intertwining with our form of cultural life: they cannot be easily grasped by concepts. In fact, expressions of aesthetic judgement show complicated roles within the culture of an historic period. According to Wittgenstein, the entire evaluative aesthetic paradigm is pluralistic: aesthetic judgments differ in different historical contexts, they assume a diverse meaning. Wittgenstein supports an aesthetic paradigm characterized by historical pluralism or contextualism. Within this perspective aesthetic concepts,—i.e. beauty, ugliness, art, aesthetic appreciation, and so on—are made and shaped inside a culture and a period in human history: they change as societies change. Shusterman has in fact affirmed that "our aesthetic concepts are inextricably bound up in our form of life, in ways of living which change over history through social, technical and even theoretical developments" (1986: 99). To understand and describe an aesthetic language game it seems necessary to acquire what Johannessen defines as the "forms of behaviour relevant to art" (2004: 28). Adopting aesthetic practices is fundamental in order to develop aesthetic sensitivity within the aesthetic field.

Thereby, Wittgenstein's description of the aesthetic field might to lead us to a dead-end. Aesthetic judgements—and, consequently aesthetic rules, concepts, meanings and patterns of reasoning—cannot

always be clear to us. Indeed, if we accept that our aesthetic concepts have a plurality of uses, as well as variable, context-dependent conditions of application, to determine an aesthetic concept would mean to describe the entire field of our aesthetic judgments, which in turn would mean describing an entire culture. At best, this is extremely difficult; at worse, it might be impossible to provide such detailed descriptions of cultures. Practising aesthetic evaluation will be just as difficult as to succeed in clarifying the aesthetic concepts of someone who is not familiar with a given culture, and therefore with the set of aesthetic rules in force within it. In the same way, aesthetic reasoning is therefore limited by our language games "which constitute our aesthetic form of life" (Shusterman 1986: 105).

5. *Fashion as an aesthetic form of life*

Keeping in mind what we have said so far about the aesthetic paradigms that can describe the evolution of fashion and considering the Wittgensteinian concepts we have addressed (PI, MWL and LC), it is now possible to delineate an aesthetic theory of fashion's mechanism. Our thesis—namely that fashion works as a "form of life"—involves Matteucci's claims on the possibility of comparing fashion to Wollheim's aesthetic theory. In this sense, fashion can be a "form of life" within a Wittgensteinian perspective. In order to demonstrate this thesis, we must apply the concept of "form of life" to fashion and see if and how this theory works.

First of all, we can note that dressing ourselves involves a set of actions that are (or are not) in agreement with aesthetic rules. That said, these rules are not normative impositions which impose us a way to dress. As Wittgenstein saw, the understanding of rules can never be exhausted by the process of rule-formulation; similarly, rules in fashion are not verbal or written fixed expressions of an *impositio* and the teaching of rules is not a mere explanation of the rule itself. The rules in aesthetics have blurred boundaries.[9] They are not eternal, since they are tied to the context in which they are to be followed as well as to the historical period in which they develop. Indeed, the aesthetic rules that govern the activity of dressing follow a temporal cyclic evolution immanent to fashion itself. Moreover, as Appelqvist affirms "the rules can be changed and abandoned as we go along" (2019: 988): the nature of grammatical rules is arbitrary (cf. MWL 7:2).

This is also the reason why it is rather hard to set an example of aesthetic rule in fashion once and for all. For aesthetic rules do not exist as platonic entities that determine how to dress properly for a specific occasion; rather, they are formed by abstraction from our daily

[9] Along with the concept of language game, in which the rules are employed in order to play (Cf. PI §71) and the correlated meaning.

practice[10]. Thus, in following a rule, one does follow an aesthetic ideal: aesthetic rules in fashion lie "in a certain consistency of our reactions and preferences, and [they] change over time (Schroeder 2020: 13). In this sense, dressing everyday according to aesthetic rules recalls Wittgenstein's concept of "following a rule": the act of dressing is a practice, indeed an action that we perform regularly, i.e. in accordance to certain rules or norms.

The notion of 'rule' is tied with what Wittgenstein called "language game."[11] Applying Wittgenstein's view to fashion results in the realization that, in dressing up for different occasions and contexts, we play different aesthetic language games with different aesthetic rules.[12] The analogy with games lets us see that, as the rules of a game change depending on the game we play, aesthetic rules might vary from one aesthetic context to another too (cf. MWL 8:87). Consequently, the meaning of a certain dress can change depending upon the context and the aesthetic rules that govern it. In fact, the meaning of clothes is dependent on the "use" we make of them in different ordinary contexts. Just as the meaning of a word can change according to its use and the context of its employment of, in the same way a garment worn in a certain way or in a certain place and time can acquire different meanings. Thus, the meaning and the sense of a dress are, to a certain extent, fluid.

The assemblage of rules—by which we can play different language games—and the correlated meanings form what Wittgenstein called "grammar". In my proposed analogy, this means that in the ordinary language of dressing we appeal to what I have labeled "the grammar of dressing". Similarly to Wittgenstein's linguistic perspective, it is possible to grasp intuitively the grammar of dressing through "a synoptic view of the grammatical system as a whole" (Appelqvist 2019: 989).[13] This way of looking at grammar does not allow to find a "conceptually determinable foundation for those norms" (Appelqvist 2019: 989): it is an expression of a 'intuitive' method to grasp meanings. Nonetheless, this does not imply that aesthetic rules cannot be defined as normative rules, as long as we consider them as part of a grammatical framework which is not eternal. Indeed, the aesthetic rules "are constitutive of the system themselves, given in the actual practices of language use" (Appelqvist 2017: 138): they are funded in the ordinary context of our routine by the repetitive action that we make when we get dressed.

To those who do not know their meaning, the rules of fashion are explained in the same way the meaning of a certain word or proposition

[10] When I say "This isn't the colour I mean; it's too cold", I don't hallucinate the colour I mean to find what ideal we're directed to, you must look at what we do: the ideal is the tendency of people who create such a thing (MWL, 9:22).

[11] Cf. section 2.

[12] One language game in fashion can be defined as "dressing to go to work" or "dressing to go to a wedding" and so on.

[13] "Taking something in as a whole at a glance" (MWL 8:59).

can be explained to someone who does not speak a certain language. In fact, precise training is required to dress properly. Partly, this training is also linguistic, for if someone is not able to understand the meaning of some clothes, then we can teach them to use fashion-related words through examples and by practice.[14] However, at the ultimate level, fashion requires a precise training in which what matters the most are the influences given by "expressions of agreement, rejection, expectation, encouragement" (cf. PI, § 208).

With exercise and training one can access the "grammar of dressing" (i.e. the set of rules on how to dress) thanks to which it is possible to learn, on the one hand, to apply the rules in the right context, and on the other hand to acquire a competent judgment on fashion-related matters.[15] This way, we can become sensitive to the rules that govern the phenomenon of fashion. And the more we become sensitive to these rules, the more we will be likely to become experts in the field of fashion. By becoming familiar with the "grammar of dressing" it is also possible to create interpretative spaces of fashion, which contribute to creating new rules of and for fashion. The fashion expert, he or she who understands fashion, is the yardstick with which to compare oneself when one is trained in fashion.

Fashion is also a source of aesthetic reactions: a shorter or longer dress can cause in us an uncomfortable reaction that can be expressed in a sign of disapproval (verbal, gestural or facial expression), as well as reactions of appreciation.[16] The frequent use of the garment denotes the pleasure one feels towards the dress, vice versa, not wearing a dress denotes dissatisfaction. In this sense, the use of a garment can express both the meaning of a dress in a certain context and the pleasure and tastes we have in aesthetic terms for certain garments. Furthermore, we can express in fashion aesthetic judgments based on aesthetic criteria of correctness. In fact, when we make an aesthetic judgment in fashion, we refer to a set of more or less evident rules, indicating the correctness (or not) of certain items of clothing or accessories.[17] But how do we know when a fashion's match is aesthetically

[14] Expressions such as "I love your skirt" or "This jacket suits you better than that one" or "This colour doesn't suit you" might be a reference to the linguistic training. Moreover, approval or disapproval, like or dislike, uncertainty and indecision can be expressed also through gestures: smiling, raising the eyebrows, rolling eyes and so on.

[15] Sensibility to the grammar of dressing is shown in our ability to discern which garment is best for each occasion (such as a white shirt for a job interview or a long dress for a wedding). Once one will acquire sensibility to the rules and context, one will be more competent in giving fashion judgments, which are expressed with advices and suggestions.

[16] A clear example is given in LC 1:13.

[17] One can express an aesthetic judgment towards a garment in every context. For example, if I'm wearing a suit or a *tailleur* for my first day at work in a formal environment, I can receive aesthetic judgments expressing correctness for my choice. Otherwise, if I'm wearing a cyan suit with amber shoes, I can receive aesthetic

correct or wrong? Here, close to the aesthetic concept of correctness, we encounter another central Wittgensteinian notion, namely the notion of the "clicking".

A "click" might be configured as a perception of correctness, which, applied to fashion, allows its functioning and diffusion. In fact, since the rules of fashion are conceptually difficult to grasp, the parameter of fashion-related judgments would also be difficult to understand if the "clicking" did not come into play. The "click" is an indication of correctness and, therefore, it is nothing more than a last proof of the correct to follow a certain rule. And it is thanks to this correctness index that we can better identify the rules that fashion dictates from time to time (cf. LC 2:8).

There are indeed some problems with this aesthetic paradigm of fashion. The first one is that the set of rules to which we appeal daily in dressing and within which aesthetic judgments are developed are almost never explicit and clear. It is very difficult (if not practically impossible) to draw up an exhaustive list of rules according to which to dress in everyday life and which are able to provide us with a stable criterion of aesthetic judgment. This difficulty emerges both at a broad level, namely the level of the basic rules of everyday dressing, and at a more specific level, dealing with the seasonality of fashion.

Since understanding and describing the rules that govern fashion is almost as difficult as defining the use of expressions of aesthetic judgment. To have a chance to succeed in understanding aesthetic judgements, you need to be familiar with the kind of aesthetic rules that govern fashion. However, to be familiar with these types of rules does not amount necessarily to be experts in the field of fashion, e.g. a stylist. It just means we can express aesthetic judgments concerning fashion, which, like the rules of fashion itself, are intimately connected with the culture of a certain historical period.

The solution to this problem could be to look at our analogy with language games: dressing up in fashion, or dressing for a specific occasion (a wedding, a theme party, etc.) are nothing more than different language games. And only by playing one game rather than another is it possible to understand the rules that guide it: in the same way it is therefore possible to understand the rules and aesthetic reactions that guide fashion. In fact, the correctness of a certain dress for a given context only emerges when two people play the same language game and can therefore assess what is right or wrong in clothing. By sharing the same language game of fashion, it is possible to grasp the rules that determine it. In this sense, we can define fashion as a particular "form of life".

judgements expressing the incorrectness of my dressing choice. In the latter case, the colour combination of cyan and amber can be followed by discomfort or disgust by the observers.

6. *Conclusions*

To conclude, considering what we have discussed so far, we can say that the combination of Wittgenstein's aesthetic-linguistic paradigm with fashion works: in fact, the meaning of a dress can change depending on the context; the rules that fashion follows are not eternal, and therefore reconcile with the properties of cyclicality and the ephemeral being of fashion. Thus, the meaning of a dress will also change as the rules concerning how to dress change. This set of rules constitutes a grammar proper to fashion, or a "grammar of dressing." Further, the practice of following a rule is consolidated thanks to a mimetic training trough which it is possible to acquire sensitivity to the rules, and thereby to becoming experts with respect to the rule. Since these rules are not eternal, it is possible to modify some of them, giving space to the need for differentiation and expression of one's identity, while at the same time not disregarding criteria of correctness or incorrectness of the way of dressing.

We can therefore say that fashion is a constellation of aesthetic language games—interpreted as sets of linguistic and cultural practices that constantly intertwine—or indeed an aesthetic language with a grammar of its own. Indeed, a form of life organizes the set of human practices in cultural and historical communities, and fashion could be one of these historical and cultural practices, which however structures or organizes itself according to its own, time-bound rules. Consider again Simmel's concept of fashion: the trickle-down theory was a model of nineteenth-century fashion that, however, can no longer be said to be valid in our times. The society of the time displayed a hierarchical structure that defined the value of fashion in a different way than today. As a result, fashion was closely linked to social and economic values of the time: it was a symbol of unequivocal social status. The relationships between social classes were rigidly vertical with respect to our society, in which, instead, there are opposite horizontal forces that guide the economic criteria of diffusion of fashion. The theoretical model of fashion outlined by Simmel was related to the *Lebensform* in which he himself lived. Simmel had, on the one hand, the great merit of discovering one of the key principles of the mechanism of fashion, namely the imitation-differentiation dialectic. On the other hand, its theoretical model cannot perfectly describe the trend of fashion today as we do not share the same form of life.

As a last test of the juxtaposition of the concepts of language game and fashionable life-form, a brief thought experiment might be helpful. Imagine being taken, without warning from the year 2019 to the year 1860, during the Victorian era on a busy street in London at the time. The first thing that passers-by intuitively would notice is how you are dressed: the substantial differences between the garments of the time (with lace, large skirts and showy hats with feathers) and today's way of dressing. Likely, even gestures and movements will be completely

different between you and nineteenth-century Londoners. In addition to this, even good taste will have completely different standards and criteria: therefore, the aesthetic judgment of nineteenth-century Londoners towards your way of dressing will, most probably, be based on the observation that your clothes are not the correct ones.

A form of life, in fact, can create shared horizons in which mutual understanding is possible and in which a sense of belonging to a sociocultural community is formed. Fashion as a form of life, in the same way, draws horizons of taste shared by the community in which to recognize oneself aesthetically. Thus, fashion as an aesthetic practice gives life to various forms of practices concerning good taste, outlining the rules that contribute to the 'grammar of dressing' of one's time: it forms a common aesthetic sense in which to move in the daily context of life. In this sense, we can say that fashion is an aesthetic form of life.

References

List of Abbreviations

MWL *Wittgenstein: Lectures, Cambridge 1930–1933* from the Notes of G. E. Moore. D. Stern, B. Rogers, and G. Citron (eds.). Cambridge: Cambridge University Press.

PI 1958 *Philosophical Investigations / Philosophische Untersuchungen*. Transl. by G. E. M. Anscombe and R. Rhees. Second edition. Oxford: Basil Blackwell.

LC 1966, *Lectures and Conversations on Aesthetics, Psychology and Religious belief*: From the Notes of Yorick Smythies, Rush Rhees, and James Taylor. C. Barrett (ed.). Berkeley: University of California Press.

Appelqvist, H. 2017. "What kind of Normativity is the Normativity of Grammar?" *Metaphilosophy* 48 (1–2): 123–145.

Appelqvist, H. 2019. "Wittgenstein: Lectures, Cambridge 1930–1933, From the Notes of G. E. Moore, edited by David G. Stern, Brian Rogers, and Gabriel Citron Wittgenstein's Whewell's Court Lectures: Cambridge, 1938–1941, From the Notes by Yorick Smythies, edited, introduced, and annotated by Volker A. Munz and Bernhard Ritter." *Mind* 128 (511): 984–993.

Borutti, S. 1993. *Per un'Etica del Discorso Antropologico*. Milano: Guerrini e Associati.

Buckley, C. and Clark, H. 2012. "Conceptualizing Fashion in Everyday Lives." *Design Issues* 28 (4): 18–28.

Dewey, J. 1934. *Art as Experience*. New York: Wideview/Perigee.

Glock, H-J. 1996. A *Wittgenstein Dictionary*. Blackwell References.

Johannessen, K. S. 2004. *Wittgenstein and the Aesthetic Domain*. In P. B. Lewis (ed.). *Wittgenstein, Aesthetic and Philosophy*. London: Ashgate Publishing Limited: 11–37.

Kawamura, Y. 2005. *Fashion-ology. An Introduction to Fashion Studies*. Oxford–New York: Berg.

Matteucci, G. 2016. "The Aesthetics as a Matter of Practice: Form of Life in Everydayness and Art." *Compendre* 18 (2): 9–28.

Matteucci, G. 2019. "Fashion." In *International Lexicon of Aesthetics*, Autumn Edition, URL = https://lexicon.mimesisjournals.com/archive/2019/autumn/Fashion.pdf, DOI: 10.7413/18258630066.

Matteucci, G. and Marino, S. 2017. *Philosophical Perspectives on Fashion*. London: Bloomsbury.

Novitz, D. 2004. *Rules, Creativity and Pictures: Wittgenstein's Lectures on Aesthetics*. In P. B. Lewis (ed.). *Wittgenstein, Aesthetic and Philosophy*. London: Ashgate Publishing Limited: 55–73.

Iannilli, G. L. 2017. "How can Everyday Aesthetics meet fashion?" *Studi di Estetica* 4 (1): 229–246.

Ratiu, D. E. 2013. "Remapping the Realm of Aesthetics: on Recent Controversies about the Aesthetic and Aesthetic Experience in Everyday Life." *Estetika: The Central European Journal of Aesthetics* 1: 3–26.

Saito, Y. 2019. "Aesthetics of the Everyday." In E. N. Zalta (ed.). *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), URL = https://plato.stanford.edu/archives/win2019/entries/aesthetics-of-everyday/.

Schroeder, S. 2020. "The Emergence of Wittgenstein's Views on Aesthetics in the 1933 Lectures." *Estetika: The European Journal of Aesthetics* LVII/XIII (1): 5–14.

Simmel, G. 1911. *Die Mode*. In *Philosophische Kultur. Gesammelte Essais*, 29–64. Leipzig: Verlag, trad. it. di Curcio, A., *La Moda*. Milano–Udine: Mimesis Edizioni.

Sheringham, M. 2006. *Everyday Life. Theories and Practices from Surrealism to the Present*. Oxford: Oxford University Press.

Shusterman, R. 1986. "Wittgenstein and Critical Reasoning." *Philosophy and Phenomenological Research* 47 (1): 91–110.

Wollheim, R. 1968. *Art and its Objects*. New York: Harpers & Row.

Croatian Journal of Philosophy
Vol. XXI, No. 61, 2021
https://doi.org/10.52685/cjp.21.1.5
Received: August 12, 2020
Accepted: December 10, 2020

Davies and Levinson on the Musical Expression of Emotion: What's the Problem?

DAVID COLLINS
McGill University, Montreal, Canada

Stephen Davies and Jerrold Levinson have each offered accounts of how music can express emotions. Davies's 'Appearance Emotionalism' holds that music can be expressive of emotion due to a resemblance between its dynamic properties and those of human behaviour typical of people feeling that emotion, while Levinson's 'Hypothetical Emotionalism' contends that a piece is expressive when it can be heard as the expression of the emotion of a hypothetical agent or imagined persona. These have been framed as opposing positions but I show that, on one understanding of 'expressing' which they seem to share, each entails the other and so there is no real debate between them. However, Levinson's account can be read according to another—and arguably more philosophically interesting—understanding of 'expressing' whereas Davies's account cannot as easily be so read. I argue that this reading of Hypothetical Emotionalism can account for much of our talk about music in terms of emotions but must answer another question—viz., how composers or performers can express emotions through music—to explain this relation between music and emotion. I suggest that this question can be answered by drawing on R. G. Collingwood's theory of artistic expression.

Keywords: Musical expression; Stephen Davies; appearance emotionalism; Jerrold Levinson; hypothetical emotionalism; R. G. Collingwood.

1. *Introduction*

It is common to speak of music in terms of emotions—to say that a song is sad, joyful, sombre, serene, etc.—and to describe pieces of music as expressing emotions, e.g., to characterize Beethoven's "Kreutzer

Sonata" as expressing agitation and pent-up energy, Vivaldi's "Spring" as expressing exuberant playfulness, and so forth. However, explaining just what is meant by such turns of phrase, and how a piece of music can be, say, sad, or how a performance of that piece can express sadness (or whatever other emotion), is a tricky business. On the one hand, musical compositions and performances are not subjects that can feel emotions, let alone externalize or express them; on the other, such talk does seem genuinely to describe a quality that we perceive or experience as somehow belonging to or being 'in' the music.

While many philosophers have endeavoured to explain how music and emotion are related and to account for talk of music expressing emotions,[1] this paper focuses on two prominent contemporary theories of musical expression: Stephen Davies's Appearance Emotionalism (AE) and Jerrold Levinson's Hypothetical Emotionalism (HE). Davies and Levinson have advanced their theories in multiple places,[2] but my focus will be on their 2006 papers "Artistic Expression and the Hard Case of Pure Music" and "Musical Expressiveness as Hearability-as-expression," respectively, since these contain the most fully worked-out and defended versions of their positions and since they explicitly frame their positions as standing in opposition to one another. This last point is important, as part of my aim is to question the view that these positions are opposing sides in a debate. I argue that, on a commonly held understanding of 'expression' that seems to be accepted by both Davies and Levinson, each position entails the other and so they are extensionally equivalent with matching truth conditions: that is, in order for one of these accounts of musical expression to be true, the other, I will argue, must also be true.

If I am right about these positions being compatible and even extensionally equivalent, one might wonder what philosophical problem is at stake in the purported debate. Although I argue that this compatibility obtains on one understanding of 'expression' and its distinction from, and relation to, what could be called 'expressiveness,' there is another understanding of expression according to which HE can plausibly be read. I argue that when HE is read in this way it no longer presupposes or is entailed by AE, making the two positions genuine alternatives. Moreover, this reading leads to HE being of greater philosophical interest and makes it better able than AE to account for nuances in our sometimes-ambiguous language surrounding expression in the arts. The real (or at least more philosophically interesting) problem in the debate on musical expression, I contend, involves the second sense of 'expression' and the philosophical problems that it raises—and which HE, but not AE, goes some way towards solving. These problems can be solved, I suggest, by looking to R. G. Collingwood's answer to a crucial question about this second sense of 'expression' and how it applies

[1] See Matravers 2007, Matravers 2010, and Kania 2017: §3, for overviews.

[2] See especially Davies 1994, Davies 1999, Levinson 1990, and Levinson 1996.

in the case of music, where HE raises but does not itself answer this question.

My discussion and argument proceeds as follows. In section 2 I outline Davies's AE and Levinson's HE, along with arguments against each position and their responses. In section 3 I argue that AE and HE are not only compatible but equivalent on an understanding of 'expressiveness' and 'expression' that is plausibly at play in the arguments for both, but that a disagreement arises on another understanding of 'expression' with which HE is more compatible albeit incomplete until a further question is answered. Section 4 summarizes Collingwood's theory of expression and shows how it can answer this question, and thus how it can complement HE so as to better explain musical expression. I conclude by considering the strengths of HE when it is supplemented with the Collingwoodian notion of expression in its primary sense.

2. *Davies's and Levinson's accounts of musical expression*

In his review of philosophical discussions of musical expression, Derek Matravers identifies four related but distinct questions of interest: "1. What is it about [a piece of] music that causes us to hear it as expressive? 2. What does 'the music is sad' mean? 3. What is it to hear music as expressive?" and "4. What is the connection between the expressive qualities of music and its value?" (Matravers 2007: 374). Since Davies and Levinson do not address the fourth question in the papers I am discussing, I shall bracket it for now. In regards to the second question, one thing that 'the music is sad' cannot coherently mean is that an occurrent emotion or psychological state is attributed to the music itself, since pieces of music are not the kinds of things that can have or be in such states. It would seem that the statement is metaphorical (or perhaps 'elliptical,' or a literal 'secondary' use of emotion terms; see Davies 2006: 183 and Young 2014: 3–5) and really attributes a property to the music that is related in some way to sadness as a felt state, i.e., in its primary sense. The question becomes: what is this property and how is it related to ordinary human sadness?

Davies and Levinson agree that an acceptable answer must centre on a property of the music rather than the composer, performer, or listener, and one that is heard or apprehended directly *in* the music rather than one that is inferred based on other of the music's properties,[3]

[3] Levinson calls these conditions the "externality requirement" and the "immediacy requirement", respectively (see Levinson 1996: 91–2). Davies's insistence that expressiveness is an "objective" property of some pieces of music (Davies 2006: 182) corresponds to Levinson's externality requirement—although, importantly, Davies clarifies that 'objective' need not mean "independent-of-human-experience", as he plausibly takes expressiveness to be a response-dependent property. Davies also criticizes accounts of expressiveness that involve "a mediated process of abstract symbolization or indirect representation" as failing to explain "the phenomenal

which rules out accounts that explain music's relation to emotion in terms of the arousal of emotion in listeners or the communication of a musician's occurrent emotion. This is not to deny that music *can* arouse emotions or manifest what a composer or performer feels but that these phenomena exhaust music's connection to emotion, from which it follows that expressiveness is a distinct—though possibly related—phenomenon. This seems right, since one can recognize that a piece of music is sad, angry, calm, etc. without feeling this way oneself, and a composer or musician can write or play an expressive piece without needing to feel the emotion it expresses when composing or performing it; as James Young notes, "Mozart … was distressed when writing some of his sunniest music" (Young 2014: 6).

Davies and Levinson also agree in limiting their discussion to so-called 'pure' music—i.e., instrumental music without an accompanying text such as lyrics, program notes, or a descriptive title that could evoke or suggest certain emotions such as "Ode to Joy" or "Morning Mood"—and to exclude purely conventional associations, since any expressiveness in these cases would not necessarily come from the music itself but the interaction of the instrumental music with the lyrics, text, or conventions (cf. Matravers 2007: 373–74). How, then, do their accounts of musical expressiveness differ?

2.1 *Appearance Emotionalism (AE)*

Davies explains the relation between a piece of music and the emotion of which it is expressive as one of resemblance, with properties of the music resembling expressive human behaviour, i.e., behaviours typical of people who feel that emotion. In this Davies follows Peter Kivy (1989), who understands expressiveness as a matter of resembling, or sharing what he called a 'contour' with, the way someone who feels a certain emotion will typically look or sound, or perhaps with the phenomenological character of that emotion.[4] For example, the faces of basset hounds and St. Bernards have a sad look due to their resemblance to facial expressions typical of sad humans—large eyes, a drooping mouth, slack muscles, etc.—, and weeping willow trees are said to be reminiscent or expressive of sadness in virtue of the resemblance between the drooping, downward turn of their branches and the slump or droop in posture of sad people, or perhaps the feelings of heaviness and being weighed down that are part of the phenomenology of sadness. Whereas Kivy and other resemblance theorists look to facial or vocal expressions or the phenomenology typical of those feeling an emotion, Davies concentrates on the dynamic character of human behaviours such as "gait, attitude, air, carriage, posture, and comportment"

vivacity with which we experience expressiveness in music" (Davies 2006: 184), where this corresponds to Levinson's immediacy requirement.

[4] For the view that expressiveness is a matter of resemblance to the phenomenological structure or felt character of an emotion, see Budd 1995.

(Davies 2006: 182). For example, people who feel sad typically walk with a heavy gait and move slowly, slump or droop in their posture, have a downcast gaze, etc., where these behaviours are constitutive of someone having 'a sad look,' whether or not they feel sad 'on the inside.' Finding the notion of an emotion having a fixed phenomenological profile to be implausible or imprecise, Davies focuses on these features since he takes them to most closely resemble an element of music, viz., its "temporally unfolding dynamic structure" (Davies 2006: 181), insofar as both involve a perception of movement, and since he sees the cross-modal resemblance between seeing—or perhaps kinaesthetically feeling—a certain sort of bodily movement and hearing the dynamic structure of music to be stronger than the resemblance between, e.g., the way an object such as a face looks and the way a song sounds.[5]

Davies's position, which he calls Appearance Emotionalism, can be stated in formal terms as: *(AE) a piece of music P is expressive of an emotion E if and only if the dynamic structure of P perceptually resembles the dynamic structure of the types of behaviour characteristic of a person who feels E.* For a piece of music to be expressive of sadness, then, is for its dynamic audible properties to resemble the "sad figure" cut by "someone who is stooped over, dragging, faltering, subdued, and slow in his movements," e.g., by the piece being "slow, quiet, with heavy or thick harmonic bass textures, with underlying patterns of unresolved tension, with dark timbres, and a recurrently downward impetus" (Davies 2006: 182). Likewise, a piece with a quick tempo, a lightness of tone or timbre, rising notes or a structural pattern that gives a sense of upward movement, harmonic resolution, etc., would be expressive of a positive emotion such as joy insofar as its dynamic structure resembled that of the behaviour characteristic of joyful people, such as an ease and lack of hesitation, lifting one's gaze upwards, dancing, skipping, or walking with a bounce in one's step, feeling free or uplifted, etc.

Davies's argument in favour of AE primarily takes the form of replies to objections, where these replies help to clarify aspects of AE so as to avoid confusions and to show how AE can account for certain aspects commonly attributed to music that might be thought to pose a challenge, such as the worry that AE would restrict the range of emotions of which music can be expressive in a way that would fail to capture the expressive nuance and subtlety that is often claimed of music.[6] The closest Davies comes to giving a positive argument in support of AE is noting that it "provides … for the phenomenal vivacity with which we experience expressiveness in music" (Davies 2006: 184), i.e., our sense of directly apprehending an emotional quality *in* the mu-

[5] While he doesn't deny that music can resemble expressive human vocalizations, he argues that the resemblance here is also mainly to do with the dynamic structure of vocal sounds, i.e., their "articulation, pitch, intensity, and periodicity of phrase lengths and shapes," rather than their timbre or inflection (Davies 2006: 181).

[6] These objections and replies are found in Davies 2006: 182–87, with the worry about the restricted range of expression being discussed on 185.

sic as we hear it, in ways that other accounts do not. James Young has recently provided further support for AE by citing empirical evidence that people experience certain dynamic structures or 'contours' of movement as being shared by instrumental music and human bodily motion, with these structures being commonly associated in each case, and often cross-culturally, with the same general emotion types (Young 2014: 19–24).

Against AE, Levinson raises the objection that talk of a resemblance between the aural appearance of a piece of music and the visual appearance of some type of human behaviour is too vague to ground a theory of expressiveness, since everything may be said to resemble, or be similar to, everything else to some degree. "The issue," he contends, "then becomes one of *how* similar such an appearance must be to one presented by human behaviour in order to constitute" the music's expressiveness of the emotion associated with that behaviour (Levinson 2006: 197). This objection is weak, since it is not clear why AE must specify a precise degree of resemblance that must be met, or at least not one that could be assessed independently of listeners' responses to music. It might appear circular to say that the music must be *similar enough* to some human behaviour to the degree that it is reminiscent of, or heard as, this behaviour; however, since this property is agreed to be response-dependent, the fact that it can only be established in reference to a listener's experience should not be seen as a problem.

There is a stronger objection to be made, however, based on a distinction Levinson makes between something being expressive of an emotion and something merely having "*an emotional quality*, in virtue of suggesting an emotion through its appearance" (Levinson 2006: 201). Since, as Levinson also notes, the notion of expressiveness depends on the concept of expression, and expression is generally agreed to be primarily a matter of a person or agent's 'inner states' being made manifest through outward appearances (Levinson 2006: 192),[7] the worry is that AE might only give a theory of what it is for music to *suggest* emotions and not of how music can be expressive of them, since a theory of the latter must make some reference to the primary sense of expression. This is where Levinson's theory comes in.

2.2 *Hypothetical Emotionalism (HE)*

Insisting that talk of something expressing or being expressive *of* an emotion is parasitic on the primary sense of 'expression,' according to which only psychological states can be expressed and hence only psychological agents can literally be said to express them, Levinson argues that in order for someone to experience a piece of music *itself* as expressive it is not enough that one registers a resemblance between properties of the music and properties of expressive behaviour. Rather, one

[7] Levinson cites Tormey 1971, Vermazen 1986, and Ridley 2003 in support.

must hear the music *as if* it were itself an expression in the primary sense; in other words, the music needs to be *heard as* an expressive act. Levinson further argues that hearing a piece of music as an expressive act involves imagining an agent or persona to whom one can attribute the act. This agent or persona need not be thought to actually exist but only be imagined as hypothetical; hence Levinson calls his position Hypothetical Emotionalism, or the Hearability-as-Expression view.

Levinson states this formally as: *(HE)* "*a passage of music P is expressive of an emotion E if and only if P, in context, is readily heard, by a listener experienced in the genre in question, as an expression of E*" (Levinson 2006: 193), where elsewhere he phrases this last clause as "*as (or as if it were) a sui generis*[8] *personal expression of [E] by some (imaginatively indeterminate) individual*" (Levinson 1990: 338). The requirement that a hypothetical persona or "imaginatively indeterminate individual" be imagined when hearing a piece of music as expressive is, for Levinson, a matter of logical entailment. If expression just is the externalization of an 'internal' psychological state via intentional 'external' behaviour, e.g., gesture, then to experience a piece of music as expressive—e.g., as an expression of sadness—can only mean to experience it as a gesture that externalizes a feeling—e.g., sadness—where the notions of behaviour or gesture and an 'inner' state presuppose an agent whose state it is and who is gesturing or behaving.[9]

One advantage of HE over AE is that it is better able to account for music's capacity to express complex, cognitively-laden emotions such as hope or embarrassment along with simple emotions such as sadness or joy, where AE would seem to restrict music's expressive range to emotions of the latter sort.[10] The strongest objection, however, is that even if it is possible to imagine a persona engaging in expressive behaviour and to hear the music *as* this expression, some or most people do not in fact engage in this imaginative activity when registering music as expressive (see Davies 2006: 190). In response, Levinson maintains that the imaginary persona needs only to be imagined "in a backgrounded manner" and is "almost entirely indefinite, a sort of minimal person, characterized only by the emotion we hear it to be expressing and the musical gesture through which it does so" (Levinson 2006: 193–94),

[8] In his 2006 paper, Levinson sets aside the '*sui generis*' claim in response to worries from Matravers (1998) about the coherence of the notion of a *sui generis* mode of expression; see Levinson 2006: 194.

[9] The persona theory of expression, or something quite like it, is also found in Cochrane 2010a (see especially pp. 203–4), Cochrane 2010b, Karl & Robinson 1997, Robinson 1994, Robinson 2005 (see especially p. 320), and, with a variation, Walton 1994. Notably, it can be found in the work of music theorist Edward Cone prior to its appearance in the work of philosophers of music (see Cone 1974).

[10] A defender of AE could insist that music can only *objectively* be expressive of such simple emotions, with more complex expressions requiring more subjective input or projection from the listener. However, see Levinson 1990 for an extended example of how a piece of music can express the cognitively complex emotion (or emotion-like psychological state) of hope. See also Karl and Robinson 1997.

and hence need not manifest in any overt or complex imaginary activity on the part of listeners such as imagining a character engaged in a narrative (cf. Davies 2006: 190). Rather, all that is needed is for one to hear passages of the music *as gestures*, i.e., to hear the music not just as movement but *as action*, where this implies an agent. For example, hearing notes that rise in pitch and increase in tempo as a 'flourish' expressing joy doesn't require one to envision, e.g., a character jumping into the air, but only involves hearing the musical movement as *someone*'s act of flourishing, with this 'someone' not needing to be any more detailed than the idea of some-flourisher-in-general.

Moreover, because it is "backgrounded," not all listeners need be overtly aware of their hearing the music as the gesture of someone gesturing: it is enough that they consciously hear it as a gesture. An even stronger reply to this worry is available to Levinson, which is to note that HE, as defined, does not require every listener to actually imagine a persona when listening to a piece of music in order for it to be expressive but only requires that the piece be "readily heard as" the expression of such a persona. In other words, all that is required is that the musical movement be *disposed* to be heard as the intentional movement or gesture of a hypothetical agent, i.e., that the music be *conducive to* being a 'soundtrack,' so to speak, for an imagined character's expressive behaviour, with this not needing be realized in the experience of every listener who finds the piece expressive.[11]

3. *What's the problem?*

Since both AE and HE can be defended against objections that have been raised against them, and since both are *prima facie* plausible (or at least HE is when understood in terms of the minimal 'backgrounded' persona that Levinson takes it to require, as opposed to some more elaborate imaginative activity), we might agree with Matravers in thinking that we "have reached a[n] impasse" in the debate (Matravers 2007: 378–79). One reason for this, I suggest, is that Davies and Levinson's accounts offer answers to two different, though connected, questions: in terms of the four questions that Matravers notes can be asked about musical expression, AE is most naturally seen as primarily answering the first question, viz. "What is it about [a piece of] music that *causes* us to hear it as expressive?", with HE primarily answering the third, viz.

[11] These responses also takes care of the objection that HE puts no constraint on the details of what multiple listeners, or the same listener on multiple occasions, imagine(s) when listening to a piece of music, which is thought to pose a problem for the convergence of different people's expressive judgments of the same piece of music (Davies 2006: 190; cf. Kania 2017: §3.1). If such detailed imaginings are not required by HE for one to register a piece of music as expressive of some emotion, the worry evaporates. While different listeners *could* engage in such diverse imaginings, HE would require the imagined personas, gestures, and narrative all to involve hearing the piece as the expression of the same sort of emotion, i.e., the emotion of which it is disposed to be heard as the expression.

"What is it *to hear music as* expressive?" (Matravers 2007: 374, my emphases). And we might think, along with Collingwood, that two statements (or theories, etc.) can only conflict if they are meant to answer the same question (see Collingwood 1939: 33) and hence that there is no actual debate to be had between Davies's and Levinson's positions.

There is a risk, however, of dismissing the debate too quickly, since both positions have something to say on both of the aforementioned questions. Even if AE holds that a resemblance between certain properties of music and certain properties of expressive human behaviour *causes* us to hear it as expressive, this implies an answer to the third question—viz., that to hear music as expressive just is to perceive this resemblance—where this might still conflict with HE's answer to this question. Likewise, even if HE holds that what it is to hear music as expressive is to hear it imaginatively as the expression of a hypothetical agent in a minimal sense, an answer to the first question is implied here too—viz., that a disposition of the music to be readily heard in this way is what causes us to hear it as expressive—where this might still conflict with AE's answer to *this* question. However, I contend that the answers AE and HE give to both questions are compatible and that, on one reading of 'expression', the positions are in fact equivalent; thus, if both theories do involve this sense of 'expression', there is no real debate between them.

3.1 *Are AE and HE equivalent?*

It is worth noting that Levinson is generally sympathetic to much of what AE claims; as Andrew Kania notes, Levinson "agrees that there is an important resemblance between the contour of music expressive of an emotion and the contour of typical behavioural expressions of that emotion" (Kania 2017: §3.1; cf. Levinson 2006: 196). While accepting the "basic thrust" of AE, Levinson's main reason for rejecting it as an adequate explanation of musical expressiveness is his worry about the notion of a resemblance between two things that do not appear "precisely" alike but are only held to be similar in appearance (Levinson 2006: 196–97). However, it is not clear why a piece of music and some expressive human movement must be precisely the same in appearance for AE to work. With Levinson's "major qualms" (Levinson 2006: 206) thus diffused, what remains is Levinson's agreement with the central condition of AE.

The compatibility of HE and AE is not only a matter of Levinson accepting that expressive music will, in its dynamic properties, generally resemble or share a contour with the dynamic properties of the kind of human behaviour that would typically be expressive of the same emotion. This resemblance is *entailed by* HE's central condition, since resembling human behaviour that is expressive of some emotion is a necessary condition for a piece of music to be readily hearable-as an agent's expression of that emotion. Consider: if a piece of music is

disposed to be heard by attentive listeners as if it were an expression in the primary sense, i.e., as a persona's gesture that expresses that persona's feeling, this will involve it disposing the listener to imagine such a gesture. If so, the piece (or the properties of the piece that so dispose it) must be perceptually similar to the relevantly analogous properties of the gesture, where this gesture is an instance of the sort of characteristically expressive human behaviour that Davies makes central to AE. That this is so can be seen when Levinson writes that, on his view, expressive music is "heard as doing something … analogous to human gesturing and vocalizing *and expressive movement, in all its forms*" (Levinson 1996: 115, my emphasis).[12] This is even more explicit in Levinson's earlier claim that "qualitative similarities and structural resemblances between the sound of a passage and standard behaviours for expressing α will typically play the largest role in *bringing about* … hearability [as an expression of α]" (Levinson 1990: 339).

That AE is a necessary condition of HE might suggest that AE offers a more fundamental explanation of musical expressiveness. However, in a similar move, Levinson argues that resemblance theories such as AE presuppose and depend on some form of 'hearability-as'. "There is," he writes, "simply no independent conception of and no access to what Davies calls musical emotion-characteristics-in-appearance *apart from* satisfaction of the hearability-in-the-music-of-an-expressing-of-emotion condition vis-à-vis attuned listeners" (Levinson 2006: 197). That is to say, one can perceive—rather than, e.g., infer—an objective resemblance between a piece of music and human behaviour *only if* there is something about the music that makes it conducive to imagining this behaviour carried out by someone or other.[13] And this holds whether or not a particular listener does in fact imagine this—although it is hard to conceive what it would be for someone to *perceive* the resemblance without in some way imagining the actions that make up the behaviour to which one perceives a resemblance.

If this is correct, then AE will entail HE. That is, *if* the dynamic structure of P perceptually resembles the dynamic structure of the sort of behaviour characteristic of a person who feels E, *then* P is readily heard (in the proper context, by a suitably receptive listener, etc.) as, or as if it were, an expression of E by some imagined or hypothetical persona. But if I am also correct about HE entailing AE, it will also work the other way around: that is, *if* P is readily heard (in the proper context, by a suitably receptive listener, etc.) as, or as if it were, an expression of E by some imagined or hypothetical persona, *then* the dynamic structure of P perceptually resembles the dynamic structure

[12] Cf. Young 2014: 9, where he writes that "a work can be perceived as the expression of an emotion *only if* it resembles the expression of emotion" (my emphasis).

[13] Cf. Davies 1999: 283, where he writes "I think that we can and must *hear music as like human action already* at the stage of hearing it as presenting expressive aspects" (my emphasis).

of the sort of behaviour characteristic of a person who feels E. And if both are correct it means that each position entails the other, making them (at least extensionally) equivalent, sharing the same truth conditions: the dynamic structure of P perceptually resembles the dynamic structure of the sort of behaviour characteristic of a person who feels E *if and only if* P is readily heard (in the proper context and by a suitably receptive listener, etc.) as, or as if it were, an expression of E by some imagined or hypothetical persona. Even if we do not accept that the two positions do not conflict because they answer different questions, the (apparently different) answer that each position gives to the same question will entail the truth of the other position, so they cannot conflict for this reason. The sense that the debate between these positions is at an irresolvable impasse is, I suggest, due to there being no actual debate to be had—at least not if AE and HE are understood according to the conception of expression in its primary sense with which Davies and Levinson are likely operating (more on which below).

Levinson pre-emptively objects to this argument when, after claiming that hearability-as-expression is often based on qualitative and structural resemblances between music and typically expressive behaviour (Levinson 1990: 339, quoted above), he writes that "it would be a mistake … to simply regard such hearability as *equivalent* to the behaviour" it resembles (Ibid., original emphasis). He gives two reasons for this. One is that it would not account for other things on which he claims hearability can also be based, e.g. conventional associations. The other is that it would overlook instances where hearability is based on a resemblance "to certain natural phenomena that have long been found expressive" (Ibid., emphasis removed) rather than to human behaviour. However, neither reason successfully counts against the sort of hearability that Levinson posits and expressive human behaviour being (extensionally) equivalent. With regard to the first, expressiveness based on conventional associations will not be 'objective' enough as a property of the music to meet Levinson's externality requirement since the convention a listener might associate with the sound of the music stands outside the music itself, whereas it is part of HE that the disposition to be heard as an expression of a certain emotion is an objective response-dependent property *of the music* alone.[14] With regard to the second, if hearability-as-expression sometimes involves resemblance to a natural phenomenon rather than to human behaviour, it's not clear that the expressiveness attributed to this phenomenon would not *itself* be an instance of AE, i.e., due to a resemblance between the natural phenomenon and characteristically expressive human behaviour. Since neither reason holds good, the pre-emptive objection fails to count against my argument above for the equivalence of AE and HE.

[14] Cf. Davies 2006: 179–80, for a dismissal of conventional associations as explaining the kind of expressiveness that is in question for both AE and HE, i.e., expressiveness as an objective property of 'pure music'.

3.2 *Expression, emoting, and betrayal*

Even if AE and HE were combined to explain musical expressiveness in terms of a property or set of properties a piece of music can have in virtue of which it both resembles and is disposed to be heard as expressive human behaviour, this explanation would be incomplete. This is because both accounts 'pass the buck' insofar as their explanations of expressiveness appeal to the concept of expression in its primary sense. For AE, the expressiveness of a piece of music is explained in terms of the expressiveness of the human behaviour which it resembles—e.g., that it has 'a sad look', etc.—, where this is in turn explained by its being the sort of behaviour that typically counts as an expression of the feelings of the one who is so behaving: e.g., being stooped over and walking with a slow, heavy, dragging gait is *expressive* of sadness because moving this way typically *expresses* sadness in the primary sense. For HE, the appeal to expression is more direct: the music's express*iveness* is explained by its being readily heard as an act of expres*sion*. But what is missing from each account is an explanation of what it is to *express* an emotion through gestures, actions, etc.

While neither AE nor HE includes a worked-out theory of expression in the primary sense, both views presuppose one. Levinson is more explicit on this, writing that to express an emotion is to reveal or make manifest an 'inner' emotional state through 'outer' signs such as "countenance, posture, bearing, demeanor, actions, gestures, and modifications of voice" (Levinson 2006: 193). Davies appears to be working with the same conception when he writes of a person's behaviour "giv[ing] direct, primary expression to the person's felt emotion" rather than merely producing an expressive appearance (Davies 2006: 183). While this might seem to be a commonsensical understanding of expression, it is too broad and thus risks conflating distinct phenomena that may often go by the name 'expression' in ordinary discourse but which need to be disambiguated if we want to gain a clear and precise understanding of expression in its primary sense, i.e., in the sense that is relevant for the current discussion. Moreover, and more importantly, whether or not AE and HE are in fact equivalent depends on their being read according to a sense of 'expression' that, I argue, should *not* be taken to be its primary sense, or at least not the sense that is primary for *artistic* expression.

The worry is that understanding expression broadly as the externalization, via behaviour, of a psychological state would include behaviour or actions that unintentionally or unconsciously manifest what someone is feeling, e.g., fidgeting as a sign of restlessness or anxiety. Indeed, this applies to some of Davies's central examples of characteristically expressive behaviour: skipping as an expression of joy often occurs spontaneously rather than deliberately, and a slumped posture and slow, heavy tread are typically unconscious and not purposely adopted by sad people. It clarifies things if, following Roger Scruton, we

distinguish these types of outward manifestation as *evincing* rather than expressing emotion[15]—or perhaps it would be more precise to speak of *betraying* emotion, as Collingwood does (see Collingwood 1938: 121–23).

One might think it that this distinction is sufficient for clarity, and that the potential for conflating expression with a similar activity could be avoided by defining expression, in its primary sense, as the *intentional* or *deliberate* manifestation of a psychological or emotional state in some outward form, e.g., through gesture or behaviour. However, this would not yet solve the problem since we can also distinguish deliberately evincing an emotion from expressing it, where we can call deliberate evincing '*emoting*' to distinguish it from a non-deliberate evincing or '*betrayal*' of emotion. This distinction between expressing and emoting is made clear by Collingwood[16] in his example of an actress who "when she is acting a pathetic scene … work[s] herself up to such an extent as to weep real tears" (Collingwood 1938: 122). He notes that merely exhibiting symptoms of grief, albeit deliberately, is not sufficient for *expressing* grief if the actress does not also "make it clear to herself and her audience what the tears are about," since expression in the primary sense means "to explore [one's] own emotions: to discover emotions in [oneself] of which [one] was unaware, and, by permitting the audience to witness the discovery, enable them to make a similar discovery about themselves" (Collingwood 1938: 122).

(In the next section I defend this account of expression in its primary sense, at least with regard to artistic expression. For now, note that one does not need to accept this definition to agree that there is a difference between, e.g., acting that merely presents, deliberately, the outward symptoms of an emotion—i.e., emoting—and *expressive* acting, which need not involve much of a display of symptoms, or at least not of those symptoms that are characteristically associated with *evincing* this emotion, in order to convey or make clear a character's emotional state to the audience in a way that also conveys something *about* the state or the character.[17])

Even if we do not yet have a worked-out and defended account of the primary sense of expression, once we distinguish expressing an emotion from both emoting (i.e., deliberately evincing an emotion) and betraying an emotion (i.e., unintentionally evincing it), it becomes clear

[15] See Scruton 1974: 214–16, especially his insistence that "the aesthetic concept of expression cannot be identified with the non-aesthetic concept of natural expression (or evincing). A gesture is a natural expression of some feeling if it is a *symptom* of that feeling, *and a symptom need not be expressive*" (214, my emphasis).

[16] Cf. Green 2016, who similarly distinguishes between what he calls expressing, representing, and showing.

[17] For an example of the distinction between acting that expresses and acting that emotes, compare the performance of Stellan Skarsgård in the Norwegian film *Insomnia* (dir. Erik Skjoldbjærg, 1997) with Al Pacino's performance, in the same role, in the American remake of the same name (dir. Christopher Nolan, 2002).

that AE conflates expressing an emotion with *betraying* one, since the characteristically expressive behaviours it takes expressive music to resemble are examples of evincing an emotion, which are held to be characteristically expressive because such behaviours are typical of people who are betraying that emotion. This is why these behaviours would be chosen to be deliberately enacted by one who wanted to *emote*, e.g., sadness or joy, or why a composer or performer might choose to play music that resembled these behaviours in its dynamic character so as to *evince* that emotion through music, resulting in 'emotive music' rather than a musical expression, properly so called, of the emotion in question. It also becomes clear that HE is equivalent to AE only if HE shares this understanding of expression in its primary sense, i.e., if it takes the expression that an expressive piece of music is readily heard-as to be the imagined emoting or betraying of an emotion by a hypothetical persona. This is not an implausible way to read HE as Levinson presents it: Matravers, for one, takes Levinson to hold that "we experience [expressive] music as the intentional communication of an emotion by virtue of its manifesting signs associated with the *betrayal* of that emotion" (Matravers 2011: 11, my emphasis).

HE can be read in another way, however, and one that is consistent with some of the ways Levinson describes expressive music. This is to take HE to require us to hear a piece of music *as if it were* the expression, in the primary sense, of an emotion, rather than hearing it as resembling or imaginable-as such an expression. In other words, on this reading the music is heard as, or as if it were, itself a (musical) gesture on the part of a persona who is imagined to be making the music to express—and not merely to evince—what he or she feels. Levinson articulates this view most clearly when he takes "[a]nother stab" at formulating HE in one of his earlier papers on the topic, writing: "Music that expresses α is music that strikes us as how a person experiencing α would behaviourally express his or her α *if* persons naturally behaved 'in music'—i.e., if the physical gestures and resulting sounds involved in playing musical instruments were a natural (unlearned, unmediated) manifestation of human emotions" (Levinson 1990: 338fn5). Admittedly, this could still be read in terms of emoting or betrayal—e.g., by imagining a world in which people's bodies naturally (i.e., automatically or non-deliberately) sometimes made musical noises due to their feeling a certain emotion, similar to how our faces naturally turn red on some occasions when we feel anger or embarrassment.[18] However, we need not imagine anything so extreme in order to realize the basic point behind this idea, which is the notion of expressing emotion directly by making music, and perhaps through gestures that are involved in producing certain sounds on certain instruments—e.g., the way one

[18] Cf. Walton 1994: 56, and Budd 1995: 132, for the interesting suggestion that we might hear expressive music "as a novel bodily means of sound production" (Levinson 2006: 194).

needs to strike a drum or cymbal in order to produce a certain sound—with these gestures being heard 'in' the music.

Needless to say, when HE is read in this way it no longer entails AE but is a distinct position, since it no longer involves imagining a persona's movements that share a dynamic contour with both the music and the emotion of which it is expressive. As well as being of greater philosophical interest, this reading of HE also gives us a more plausible account of musical expression than one based on resemblance to behaviours such as gait, posture, and the rest. This is because, at the level of generality that is involved in registering a resemblance between dynamic contours, there would seem to be no way to distinguish behaviour that expresses an emotion from emotive behaviour or behaviour that betrays an emotion when discerning which behaviour the contour of a given piece resembles. Unfortunately, Levinson appears not to have pursued this idea as far as he might have due to its initially being presented in relation to his claim of music having a *sui generis* form of expression, where he backed off from this idea in the face of criticisms that were raised against it. These criticisms do not count against the idea of expressing oneself through making music, since this need not be taken as *sui generis* in the objectionable sense of being a wholly unique mode of expression, and so this initial idea of Levinson's is worth taking up again and exploring further.

Reading HE in this way gives us an explanation of musical expressiveness in terms of a disposition that some pieces have to be readily heard as the kind of music someone might make if they were expressing an emotion by making music (in that genre, with those instruments, etc.), with the imaginary persona being the hypothetical composer or performer of the music we are hearing—where this persona will coincide with the actual composer or performer in cases where a piece of music (i) actually was an expression of the composer or performer's emotion, and (ii) was heard as expressive of that emotion by a listener. Basing expressiveness on this understanding of expression in its primary sense, and moreover on the notion of expressing an emotion by making music, clarifies that the *real* philosophical problem concerning musical expressiveness lies in explaining what it is to express an emotion in and through music, as opposed to emoting or evincing an emotion by means of music. Solving this problem requires us to have some account of artistic expression in general and to know how this would apply to music as a medium, and it is here that looking to Collingwood's account of expression can help.

4. *How Collingwood can help solve the 'real problem' of musical expression*

One difficulty for any appeal to Collingwood's theory of art is that it has been frequently misconstrued in such a way as to make it seem

implausible, to the point where this misconstrual has come to be accepted by many, if not the majority, of philosophers of art. As such, it is worth showing that Collingwood's theory is more plausible than this misreading allows so as not to turn off readers who might suspect that supplementing HE with Collingwood's theory of expression might weaken the resulting account in terms of its plausibility and its explanatory value for actual instances of musical expression.[19] The misreading takes Collingwood to hold that artworks are mental objects existing solely within the minds of artists and audiences, with the tangible object or event that is commonly called the work—e.g., a physical painting or sculpture, a dancer's bodily movements through space, or in the case of music the sounds produced by instruments when played—being only contingently related to the real work of art as the means by which this work can be transmitted from the artist's mind to the minds of audience members. Admittedly, some of Collingwood's statements make it sound like this is what he is saying, especially when read outside their full context. However, Aaron Ridley and others have shown that despite these passages that seem to support the misconstrual, Collingwood does not hold this 'ideal theory,'[20] and that when he makes an apparently damning remark, e.g., that a musical work "is [a] tune in the composer's head," where "[t]he noises made by the performers, and heard by the audience, are not the music at all; they are only means by which the audience, if they listen intelligently … can reconstruct for themselves the imaginary tune that existed in the composer's head" (Collingwood 1938: 139), he is not talking about a mental as opposed to a physical object but rather is making the point that a set of sounds is not music unless these sounds are heard together as forming a tune or pattern.

What, then, is Collingwood's theory of artistic expression if not the 'ideal theory' it is so often mistaken for? In brief, he takes expression to be the articulation of the felt or qualitative dimension of an experience, where this articulation is necessarily carried out in a medium—e.g., in paint, stone, gesture, sound, language, etc.—through a non-technical process in which the artist comes to discover the end she is working towards (i.e., the completed artwork) in the process of creating it, rather than beginning with a clear conception of the work and then figuring out the means to realize it. This articulation can also be understood as a clarification of the artist's pre-conscious feelings that brings them

[19] This suspicion would be unfounded in any case since, as an anonymous reviewer noted, the metaphysics of Collingwood's theory of art, which is the part of his theory that is commonly misconstrued so as to make it seem implausible, can be separated from his account of expression. Even so, it seems worth taking a paragraph to note how this common reading is mistaken in order to avoid possible distraction or suspicion on the part of readers who are only familiar with Collingwood through this misconstrual.

[20] See Ridley 1997 for his refutation of the 'ideal theory' reading, as well as D. Davies 2008 and Wiltshire 2018 for further defences of Collingwood's theory that are compatible with a non-idealist reading. For the most influential presentation of the idealist interpretation of Collingwood (i.e., the misconstrual), see Wollheim 1972.

to conscious awareness, making them comprehensible, or graspable in thought, for the artist and her audience alike.

The process of expression begins with an indeterminate feeling at what Collingwood calls the 'psychical' level of experience. This is something like a pre-conscious, embodied registering of sensations and affects occurring at the fringes of awareness, with sensations and their affective charges together being what Collingwood calls 'feelings'. At this point, precisely *what* one feels remains indeterminate; as Collingwood writes, "When a man is said to express emotion … [at] first, he is conscious of having an emotion, but not conscious of what this emotion is. All he is conscious of is a perturbation or excitement, which he feels going on within him, but of whose nature he is ignorant. While in this state, all he can say about his emotion is: 'I feel … I don't know what I feel'" (Collingwood 1938: 109). *Expressing* this feeling involves first directing one's attention to it, which in Collingwood's terms 'raises it to consciousness'—or, we could say, shifts the focus of our attention so that the feeling is no longer on the 'fringe' of our awareness but closer to the centre. This allows one to apprehend and fix the psychical feeling before one's mind, which makes the feeling more determinate and available for thought and reflection, although at this stage the feeling is not yet clarified.

Once the feeling has begun to be consciously grasped, the artist can start to get clear on what, precisely, she feels by articulating this feeling in a medium and turning it into an 'emotion for consciousness.' The emotion is articulated when the initially pre-conscious and indeterminate 'psychical' feelings are given form in the chosen medium; we might say that the emotion is now embodied in the perceptible form, with perceiving the final form being identical with consciously apprehending the clarified emotion (cf. Sias and Bar-On 2016). Because this is a non-technical process with the artist not yet being fully aware of the nature of her emotion, she must proceed, as Ridley (2003: 222) notes, by feeling her way along, continuing to attend to the qualitative aspects of her experience as she engages and interacts with her medium. Collingwood gives an example of a painter stepping back to observe his work as it develops and making adjustments and revisions in response to what he sees, thinking as he goes "I am not satisfied with that line; let us try it this way … and this way … and this way … there! that will do!" (Collingwood 1938: 281). Likewise, a writer in the process of composing a poem will try out certain words and phrases, varying their order or substituting certain words with others, changing their punctuation and spacing, etc. until she arrives at the precise form that she apprehends as being 'just right.'

With the basic elements of Collingwood's account of expression now outlined,[21] we are able to explain what it would be to express, rather than just to evince, an emotion in and through music. This would be

[21] See Ridley 2003: 222–25 for another summary of the basics of Collingwood's theory, which Ridley contrasts with transmission and resemblance-based theories of expression, presenting it as 'expression proper.'

a matter of articulating an initially vague feeling—the qualitative or affective aspect of an experience—by producing or combining sounds in such a way so as to give this feeling audible form and thereby make it available for a listener's awareness as a determinate emotion. While there would be no single way for this articulation to proceed, two examples will clarify the sort of activity this might involve in practice.

(1) A musician has long been fascinated by a particular piece, sensing that the composition has greater aesthetic potential than he has so far been able to realize in his performances of it, although he cannot yet say precisely in what this potential consists. Rehearsing this piece for yet another performance, he experiments with different ways of playing it, making slight adjustments in certain passages. Through trying out a number of these variations, guided by instinct and a feeling that he will know what he is looking to achieve when he hears it, he starts to develop a better idea of this intuited potential by getting a feel for which variations seem on the right track and which do not. Eventually, through trial and error, he becomes aware of just what the potential was that he sensed in the piece, where his becoming aware of this is identical with his first playing the piece in a particular way and recognizing it as being what he was looking for: this new performance successfully expresses how he feels about the piece.

(2) A composer who is setting out to write a new piece without feeling particularly inspired is playing around on her piano when a certain arrangement of notes that she happened to play strike her as sounding particularly evocative—but *of what* she cannot say. As she explores ways of developing and elaborating upon this musical phrase, riffing on and modifying it so that it ends up being very different from the notes she initially played while retaining, and amplifying, the initial feeling that struck her, she becomes aware that the song she is writing expresses elements of her mood that morning, making it more clear to her how she had been feeling when she sat down to compose.

With this explanation of what it would be to express an emotion by making music, we can also say what it would be to hear a piece of music as doing this and thereby complete the non-resemblance-based understanding of HE discussed above. A piece of music is expressive of an emotion, on this account, if and only if it is readily heard by qualified listeners as (or as if it were) the result of a process of articulating an initially indistinct sensation or affect that gives it a determinate perceptible (audible) form in an arrangement of sounds in which it is embodied. Note that this does not require a piece *to be* an expression in order for it to be expressive, only that it be readily heard *as* one, i.e., that it is conducive to being imagined as the (musical) expression of a hypothetical persona. This could be a matter of imagining a composer

or performer expressing a feeling through writing or performing the piece in question in just the way it is heard, and imagining an emotional state of the composer's or performer's that corresponds to and is articulated by the piece the listener hears. This could also be a matter of a piece being conducive to being heard in fresh ways on multiple occasions, so that a listener is able to imagine that the new elements and dimensions of the piece that she 'discovers' upon a repeat listening correspond to newly clarified aspects of the feeling that a hypothetical persona is articulating for the first time via the music, perhaps identifying this persona with herself and imagining that she is articulating her own feelings through the music as she hears these new elements in it.

One worry that arises for this explanation is that it risks making the ability to register music as expressive rare, since being a 'qualified listener' would seem to require that one be familiar with music from the standpoint of a composer or performer in order to be aware of the expressive possibilities offered by the medium. Including the 'qualified listener' requirement in the definition means that some listeners might not be positioned to experience a piece of music as expressive, being limited to music merely arousing emotion or reminding them of one by evincing it in the way AE posits, although I do not think we need to worry that it would restrict the former ability to musicians and musicologists. Even if not everyone has experience with composing music or playing an instrument, nearly everyone has direct personal experience making sounds and being aware of what it feels like to make them, and moreover what it feels like to hear them as one is making them,[22] where this is plausibly sufficient to allow us to have *some* sense of what it might feel like for us to make some other sound, e.g., one on an instrument that we do not actually play, and so what we might be expressing by making that sound. This is compatible, of course, with the view that the more experienced one is in playing, composing, or studying music the more likely one will be to have a *better* sense of the expressiveness of a piece, which I think is also plausible.

Another worry about this explanation is that it is still too vague, since it does not tell us exactly what the expression of various emotions through music will involve in a way that would allow us to recognize musical expressions for what they are when we hear them, where one might think that AE succeeds in doing this by telling us that sad music typically will sound slow, heavy, downbeat, descending, etc. This objection is, I think, mistaken, since it assumes both that the emotions that artworks express are general types—i.e., what Collingwood calls "thing[s] of a certain kind"—instead of being particular feelings that are uniquely expressed by a particular work—i.e., what he calls "certain thing[s]" (Collingwood 1938: 114)—and that just what will be

[22] Cf. Collingwood's point that every speaker is also conscious of him- or herself as speaking, and so is both a speaker and the first member of his or her audience (Collingwood 1938: 247–51).

involved in expressing an emotion can be known and specified in advance of its expression. This is to treat artistic expression on the model of what Collingwood calls 'craft' rather than as art, where artworks and their expressiveness are not determinable in this way since expression just is the process of clarifying what is not yet clear.

Moreover, the fact that AE *can* specify in advance what a song that is expressive of sadness, joy, etc. will *typically* sound like shows that it assumes a craft-based or technical conception of art, whereas on Collingwood's account there is nothing that it characteristically sounds, or looks, like for a artwork to *express*, as opposed to evince, an emotion. I take the plausibility of this conclusion to be another point in favour of Collingwood's account of expression, as well as a point in favour of HE as an explanation of musical expressiveness since it can be read as compatible with Collingwood's account in a way that AE cannot.

5. *Conclusion*

I have argued above for three main points: (i) that AE and HE are compatible and extensionally equivalent theories of musical expressiveness when read according to the understanding of the primary sense of expression with which Davies and Levinson are most readily taken to be operating, and that as such there is no genuine debate to be had between them; (ii) that HE can be read in another way according to which it is not equivalent to AE and is more philosophically interesting, although incomplete; and (iii) that Collingwood's theory of expression can be used to complete HE by explaining what expressing an emotion in and through music would involve.

One virtue of looking to Collingwood for an explanation of expression in its primary sense is that his account—when *not* misconstrued as the 'ideal theory'—takes artworks to embody the emotions they express in perceptual form,[23] and so is compatible with Levinson's externality and immediacy requirements for expressiveness. Another is that HE, read in terms of hearability-as-expression-*in-music* and according to Collingwood's account of expression, can meet many of the objections commonly brought against both expression and resemblance theories of expressiveness. Some of these are discussed above; two that are not are those that Kania (2017: §3.1) identifies as the main problems for expression theories of musical expression. First, HE on this reading allows that a composer or performer might not themselves feel an emotion that the music they make is expressive of, since not every expressive piece of music needs actually to be a musical expression but needs

[23] When a work is one of the rare few that Collingwood *does* allow could be created entirely 'in the artist's head'— such as a short poem or tune that the artist does not need to write down, speak, or play in order to fully work out—it will still have a form that is imagined as quasi-perceptible (e.g., the artist imagining hearing it played, seeing the words on the page, etc.) and so in principle could be externalized and made publicly available to others.

only be readily heard *as if* it were one. Second, it allows for a composer to fail to express her emotions in a piece she writes since it does not tie a piece's expressiveness to what an artist is actually feeling, but only to its hearability as the kind of music that someone who *was* feeling this *might* plausibly make to express it. This suggests that a successful musical expres*sion* must also be expres*sive* in the sense of being readily heard by listeners as the expression that it in fact is, which is a philosophically interesting and, I think, plausible element of artistic expression—but one that should be explored in a future paper.[24]

References

Budd, M. 1995. *Values of Art: Pictures, Poetry and Music*. London: Penguin.

Cochrane, T. 2010a. "A Simulation Theory of Musical Expressivity." *Australasian Journal of Philosophy* 88: 191–207.

Cochrane, T. 2010b. "Using the Persona to Express Complex Emotions in Music." *Music Analysis* 29: 264–75.

Collingwood, R. G. 1938. *The Principles of Art*. Oxford: Clarendon Press.

Collingwood, R. G. 1939. *An Autobiography*. Oxford: Clarendon Press.

Cone, E. T. 1974. *The Composer's Voice*. Berkeley: University of California Press.

Davies, D. 2008. "Collingwood's 'Performance' Theory of Art." *British Journal of Aesthetics* 48: 162–74.

Davies, S. 1994. *Musical Meaning and Expression*. Ithaca: Cornell University Press.

Davies, S. 1999. "Response to Robert Stecker." *British Journal of Aesthetics* 39: 282–87.

Davies, S. 2006. "Artistic Expression and the Hard Case of Pure Music." In M. Kieran (ed.). *Contemporary Debates in Aesthetics and the Philosophy of Art*. Oxford: Blackwell: 179–91.

Green, M. 2016. "Expressing, Showing and Representing." In C. Abell and J. Smith (eds.). *The Expression of Emotion: Philosophical, Psychological and Legal Perspectives*. Cambridge: Cambridge University Press: 25–45.

Kania, A. 2017. "The Philosophy of Music." In E. Zalta (ed.). *The Stanford Encyclopedia of Philosophy*. Online: URL = <https://plato.stanford.edu/archives/win2017/entries/music/>.

Karl, G. and Robinson, J. 1997. "Shostakovitch's Tenth Symphony and the Musical Expression of Cognitively Complex Emotions." In J. Robinson (ed.). *Music and Meaning*. Ithaca: Cornell University Press: 54–78.

Kivy, P. 1989. *Sound Sentiment: An Essay on the Musical Emotions Including the Complete Text of The Corded Shell*. Philadelphia: Temple University Press.

Langer, S. 1953. *Feeling and Form*. New York: Scribner's.

Levinson, J. 1990. "Hope *in* The Hebrides." In J. Levinson, *Music, Art, and Metaphysics*. 2011 re-issue. Oxford: Oxford University Press: 336–75.

[24] Thanks to David Davies, Moe Touizrar, and the audience at the 2014 meeting of the Canadian Society for Aesthetics, especially James Young and John Brown, for their feedback on the ideas and arguments in this paper.

Levinson, J. 1996. "Musical Expressiveness." In J. Levinson, *The Pleasures of Aesthetics*. Ithaca Cornell University Press: 90–125.

Levinson, J. 2006. "Musical Expressiveness as Hearability-as-expression." In M. Kieran (ed.). *Contemporary Debates in Aesthetics and the Philosophy of Art*. Oxford: Blackwell: 192–204.

Matravers, D. 1998. *Art and Emotion*. Oxford: Clarendon Press.

Matravers, D. 2007. "Musical Expressiveness." *Philosophy Compass* 2/3: 373–79.

Matravers, D. 2010. "Recent Philosophical Work on the Connection between Music and the Emotions." *Music Analysis* 29: 8–18.

Ridley, A. 1997. "Not Ideal: Collingwood's Expression Theory." *The Journal of Aesthetics and Art Criticism* 55: 263–72.

Ridley, A. 2003. "Expression in Art." In J. Levinson (ed.). *The Oxford Handbook of Aesthetics*. Oxford: Oxford University Press: 211–27.

Robinson, J. 1994. "The Expression and Arousal of Emotion in Music." *The Journal of Aesthetics and Art Criticism* 52: 13–22.

Robinson, J. 2005. *Deeper than Reason: Emotion and its Role in Literature, Music, and Art*. Oxford: Oxford University Press.

Sias, J. and Bar-On, D. 2016. "Emotions and Their Expressions." In C. Abell and J. Smith (eds.). *The Expression of Emotion: Philosophical, Psychological and Legal Perspectives*. Cambridge: Cambridge University Press: 46–72.

Scruton, R. 1974. *Art and Imagination*. 1998 re-issue. South Bend: St. Augustine's Press.

Tormey, A. 1971. *The Concept of Expression*. Princeton: Princeton University Press.

Vermazen, B. 1986. "Expression as Expression." *Pacific Philosophical Quarterly* 67: 196–224.

Walton, K. 1994. "Listening with Imagination: Is Music Representational?" *The Journal of Aesthetics and Art Criticism* 52: 47–61.

Wiltshire, N. 2018. "Feeling, Emotion and Imagination: In Defence of Collingwood's Expression Theory of Art." *British Journal for the History of Philosophy* 26: 759–81.

Wollheim, R. 1972. "On an Alleged Inconsistency in Collingwood's Aesthetic." In M. Krausz (ed.). *Critical Essays on the Philosophy of R.G. Collingwood*. Oxford: Clarendon Press: 68–78.

Young, J. O. 2014. *Critique of Pure Music*. Oxford: Oxford University Press.

Croatian Journal of Philosophy
Vol. XXI, No. 61, 2021
https://doi.org/10.52685/cjp.21.1.6
Received: July 5, 2020
Accepted: November 3, 2020

Learning from Fiction to Change our Personal Narratives

ANDREW J. CORSA
Lynn University, Boca Raton, Florida, USA

Can fictional literature help us lead better lives? This essay argues that some works of literature can help us both change our personal narratives and develop new narratives that will guide our actions, enabling us to better achieve our goals. Works of literature can lead us to consider the hypothesis that we might beneficially change our future-oriented, personal narratives. As a case study, this essay considers Ben Lerner's novel, 10:04*, which focuses on humans' ability to develop new narratives, and which articulates a narrative that takes into account both everyday life and large-scale issues like the global, environmental crisis.*

Keywords: Narrative; fiction; Ben Lerner; autofiction; global warming; living option.

When I read Ben Lerner's (2014) book, *10:04: A Novel,* I felt compelled to reflect on my capacity to create new, future-oriented personal narratives for myself. In the course of my reflections on *10:04*, I developed a question that became the focus of this essay: Can fictional literature help individuals create new future-oriented narratives for themselves that they might not have created otherwise, and might those narratives help them lead better lives in which they are more successful at achieving their goals?

The first part of this essay provides a general background about personal narratives, giving a sense of what theorists such as Arran Gare and J. David Velleman have suggested about the role future-oriented personal narratives play in guiding our actions, and why changing our personal narratives might sometimes help us to more successfully pursue our goals. The second part of this essay argues that works of literature can lead us to consider changing our personal narratives in ways that we might not have considered otherwise. Building on the work of

theorists such as Catherine Z. Elgin and David Novitz, this essay argues that, as a result of reading works of literature, some readers might consider the hypothesis that adopting certain narratives, like those described by the works of literature, might help them to better achieve their goals. This essay argues that, after reflecting on certain works of literature, some readers might consider the possibility of adopting a new narrative to be a genuine, living option. While many theorists have argued that literature can have a positive impact on readers, this paper's discussion of narratives becoming living options is unique.

To provide a case study, this essay reflects on *10:04*, which focuses in part on environmental concerns like global warming. This essay argues that, after reading *10:04,* some readers might develop for themselves, and consider, the hypothesis that they should develop new personal narratives that better balance everyday life with a concern for the earth's environmental crisis. Part four examines how *10:04* can effectively lead readers to reflect on the ways in which fictional literature helps them create new personal narratives.

1. *Personal narratives*

This essay understands narratives as structured stories which people tell either themselves or others, which may be either fictional or true-to-life (cf. Goldie, 2012: 2). Narratives are structured with a beginning, a middle, and end, and reveal causal connections between many of the events they describe (Goldie 2012: 11–17; Robinson 2014: 310–312). As Grace Clare Robinson notes, while many narratives are expressed verbally, they need not be, and narrative content, which is the same as the "basic structure" of a story, can "survive translation" into different media (Robinson 2014: 311; cf. Carr 1986: 62–63). Robinson claims, for instance, that it would be possible for the narrative of Romeo and Juliet to be translated into a silent dance. Peter Goldie contends that narratives need not be publicly expressed; they can be "thought through in narrative thinking" (2012: 2–4; cf. Carr 1986: 62–63). People routinely develop narratives in their minds before they articulate them to others, and often never choose to publicly articulate their mental narratives.

Arran Gare claims that people often make sense of the world through narratives—through the stories that we tell ourselves and others (1995: 62; cf. MacIntyre 2008: 216). David Carr maintains that while some writers of literature articulate fictional narratives, humans in general are often inclined to organize the events of their own real lives into narratives (1986: 4–5; cf. Bruner 1991: 18; Kirkpatrick 2014: 60; Bruner and Kalmar 1998: 318). Many humans also routinely construct narratives about what their own futures might be like—stories about what they intend or hope to do.[1] According to both Gare and

[1] For discussions of "narrative thinking about one's future," see: Goldie 2012: Chapter 4.

Carr, we routinely construct such future-oriented narratives when we pursue "long-term undertakings" such as "writing a book, getting an education, [or] raising a child" (Carr 1986: 54 and 61; cf. Gare 2001).[2]

Humans do not wait until their undertakings are complete before constructing narratives about them. Instead, we build narratives to help us complete the undertakings while they are ongoing. Carr writes that our complex undertakings can be understood "as a process of telling ourselves stories, listening to those stories, and acting them out or living them through" (1986: 61; cf. Gare 2001; MacIntyre 2008: 212; Corsa 2018: 243–244). The narratives we tell ourselves about our futures guide our actions. As Goldie notes, sometimes when we construct narratives in our minds about actions we intend to perform, our "narrative thinking might turn into a plan do just that" (2012: 88). As J. David Velleman puts it, a person might sometimes "narrate ahead of himself and then follow a career that reflects his story" (Velleman 2005: 63).

Velleman contends that, by composing narratives about how we hope to act, we become more likely to act as we hope (Velleman 2005: 65; cf. Tollefsen and Gallagher 2017: 101–103). At the end of a day of work, you might say to yourself, "I am going home now," and construct a mental narrative in which you in fact, go home—not because you were already going to leave, but in order to prompt yourself to do so. Velleman notes that there are many cases in which expressing that you will perform an action makes you more likely to perform it. For example, he refers to a psychological study, performed by Greenwald, Carnot, Beach and Young (1987), which indicates that subjects who predict they will vote are significantly more likely to vote than those who are never asked to make a prediction (Velleman 2005: 65–66). Carr maintains that by telling ourselves future-oriented stories about how we will complete our undertakings, we can also: remain clear about how our current activities relate to our goals; clarify to ourselves what we will need to do next; better determine if we have gone off track; and better figure out if we need to change our strategies to address changing circumstances (1986: 61, 71 and 87; cf. Corsa 2018: 243–244). The stories we tell ourselves can even affect our view of the world; Jerome Bruner and David A. Kalmar contend, for example, that if we construct

[2] Unlike several theorists discussed, this essay does not argue in support of the claim that selves are constituted by narratives. While this essay's contentions are consistent with this claim and would fit well in many theories that support it, many objections to this claim are not objections to this essay's arguments. Suppose the claim that selves are constituted by narratives is false. Also suppose it is possible to have coherent self-experience without "understanding and living out our lives in storied ways" (Hutto 2016: 29). Nonetheless, this essay's arguments could still be sound. It could still be true that: people often make sense of the world through narratives; many people construct narratives about their futures; these narratives change their actions; and developing new personal narratives might sometimes help people achieve their goals. Daniel D. Hutto, who rejects strong narrativism, nonetheless maintains that people often make use of narratives in their efforts to shape their lives (Hutto 2016: 25).

personal narratives in which we are "victims," we will likely see others as more agentive than if we saw ourselves as "agents," instead (1998: 317).

The way that people think of their personal narratives, in real life, does not closely resemble the way that critics think of narratives in works of literature (cf. Lamarque 2007: 118; Schechtman 2011: 413). As Peter Lamarque notes, when critics read works of literature, they often take chance events and small details to have symbolic significance, to foreshadow future events, or to help develop themes (Lamarque 2007: 123–125). A critic might say that an event in a literary narrative, which characters view as a chance accident, needed to happen in order for the author to effectively develop a theme, a character arc, or a plot (Lamarque 2007: 124–127). Now consider real-life personal narratives. Lamarque notes that it would be a mistake to frequently attribute symbolic or thematic significance to small details or chance events in our lives (Lamarque 2007: 131; cf. Schechtman 2011: 411–413). Likewise, it would be an error to think that every chance event needed to happen, in order for our lives to take a shape they were "supposed" to have (Lamarque 2007: 131; cf. Schechtman, 2011: 411–413).

Rather, as Marya Schechtman indicates, while the way we reflect on our real-life narratives should partly resemble the way that critics think of literary narratives, it should also partially resemble the way that authors and characters view their narratives (2011: 413). In many works of literature, characters do not think of small details or chance events as having symbolic or thematic significance, or as needing to happen in order to develop a specific theme. Like characters, it would not be appropriate for us to think of the small details or chance events in our lives in these ways, either. Also like most characters, real people do not have complete control over their narratives; we cannot completely determine how other people will react to us (Schechtman 2011: 413), or change what actually happened in our past (cf. Vice 2003: 103). Yet, as Samantha Vice notes, it would be mistaken to think that we, in real life, are like characters in a stories that have already been written, and that we need to stick to certain roles or plot-lines (Vice 2003: 103). Instead, as Schechtman indicates, in some respects we need to view our lives like authors view stories that are not yet finished; we have some control over our future narratives insofar as we are responsible for our choices (2011: 413–414). Finally, like critics, we can reflect on the significance of our choices as authors of our own personal narratives, and doing so might help us to make better choices in the future.

Many of the theorists discussed above agree that the narratives we construct about ourselves have an influence on the actions we take, our perspectives, and whether we are successful at achieving our goals. Our lives shape our stories, but our stories also shape our lives (Velleman 2005: 64). It follows that sometimes it might be prudentially valuable for individuals to change the personal narratives they tell

themselves or others, in order to change their future behaviors or outlook. That way, they might more effectively reach their goals or, alternatively, strive for different goals. Gare writes that individuals who choose to "question the stories they have been socialized into" and live their lives in accordance with "alternative versions of these stories" that they have deliberately chosen for themselves "take responsibility for their lives" and "are the creative agents of culture, society and history" (2002: 97; cf. Gare 2001; Kirkpatrick 2014: 65; Carr 1986: 92–9; Bruner and Kalmar 1998: 326).

2. *Learning from literature*

We might sometimes be unable to create new narratives that would better help us achieve our goals because we are unaware of the kinds of narratives we could create. When reflecting on what options are open to us, we are often constrained by past experiences, by emotional dispositions, and by personal character (Goldie 2012: 89).

As Catherine Z. Elgin notes, people routinely engage with a huge quantity of information; objects of perception and life experiences have a vast number of properties. We inevitably have to ignore some information and some properties, focus substantially on others, and attend to only some possible patterns (Elgin 2007: 44). Elgin writes that we are often unaware of how much information we ignore and how we choose to focus our attention; rather "we adopt familiar orientations and judge by received standards" and we have "cognitive default settings that we invoke unthinkingly" (2002: 2). We tend to focus on properties, categories, and patterns based on our established habits and "time-honored methods" (2007: 46). When we determine how to act, make plans, and create future-oriented personal narratives, we rely on the information, categories and patterns on which we focus. Since we are often unaware of the information and patterns we ignore, "we are blind to the mistakes we might be making and the opportunities we might be missing" (Elgin 2007: 46). We might fail to develop better future-oriented narratives for ourselves, because we are limited by our outlooks, and we are unaware of the ways in which we could change them. Certainly, it is possible for us to meet and observe people we deem successful and admire, but without understanding the very different ways in which they categorize what they perceive—and the different ways in which they focus on and ignore information—we might sometimes not understand the personal narratives they developed for themselves,[3] or how we might develop and live according to related narratives.

Even if we do understand what sorts of future-oriented personal narratives we might practically develop for ourselves, we might not think of some of those narratives as "living," genuine options. Reflect-

[3] Compare with the different but related considerations in: Gallagher 2014: 602–603 and 606.

ing on the philosophy of William James, Alexis Dianda maintains that a living hypothesis is one that must cohere with a person's beliefs, cultural context, and history (2018: 651–652; James 1898). Catholics, but not atheists, might accept as "living" the hypothesis that it would be beneficial to create a future-oriented personal narrative according to which they attend Catholic mass regularly for the rest of their lives. Likewise, Goldie notes that our character traits constrain what options and future-oriented narratives we consider "living"; after all, a "considerate person wouldn't *dream* of leaving someone in the lurch" and "the punctual person just wouldn't *think* of dawdling" (2012: 89). Additionally, a person might choose not to develop or try to live according to certain future-oriented narratives, believing that they could not possibly live according them, even when they really could. There is a risk that people might sometimes reject personal narratives that would actually help them to better achieve their goals or be the kind of people they would most like to be.

Some works of fictional literature, however, can help individuals to better recognize what kinds of personal narratives they might develop for themselves, and to recognize those narratives as genuine, living options for them. This essay does not contend that readers routinely use fictional literature this way, or that all works of literature could be effectively used like this. Rather, it maintains that, after reading *some* works of fictional literature, *some* readers might reflect on the hypothesis that developing a different personal narrative might help them to live better.[4] Some works of literature might also help readers to recognize that the hypothesis on which they reflect is living and possible.[5]

Obviously, fictional literature includes many statements that are not true of the real world. Even when literature includes true statements, readers often should not consider them true of the real world without additional evidence. For example, as Elgin notes, *Moby Dick* includes truths about what methods are effective for harpooning a whale. Yet, without additional evidence, a reader could not know when the book's author has taken liberties with the truth, could not know which of its claims are true and which false, and ought not consider any of its claims about whaling true of the real world without additional evidence (Elgin 2007: 43).

Novitz proposes that we can treat fictional literature as offering hypotheses about what the world is like, which readers can then test out in reality, and either confirm or reject based on life experience, or based on whether each hypothesis "coheres with their established beliefs" (Novitz 1984: 56–57; cf. Vidmar 2013: 178–179; Elgin 2007: 53).

[4] While Nussbaum does not discuss hypotheses or narratives in these ways, this essay was influenced by her arguments about how works of literature help us to grapple with the question: "How should one live?" See: Nussbaum 1990. Nussbaum frames this question on p. 50.

[5] For discussions of how William James thinks people can transition from seeing a hypothesis as dead to seeing at as living: Dianda 2018: 656–659.

In some cases, a reader might study the real world, in order to determine whether a fiction's hypothesis is true. Or, as David Davies indicates, if a work of fictional literature indicates a pattern underlying the experiences of its characters, we might "test" the hypothesis that this pattern also belongs to real-life experiences by reflecting on whether we "feel" that the pattern is actually apparent in the real world; we might trust such feelings because we take them to be based on "tacit or unarticulated knowledge" of the world we already had prior to reading (Davies 2007: 44; cf. Vidmar 2013: 179 and 190; Kivy 2006: 107–108). Peter Kivy further contends that works of literature can sometimes help readers to see certain hypotheses as "living" rather than "dead" (2006: 101–102; cf. Vidmar 2013: 190). What this essay adds is that, after reading some works of fictional literature, some readers might consider hypotheses about personal narratives they might want to adopt and realize that these hypotheses are living.

How do fictions do this? They present us with narratives very different from our own, and after reading them, we might sometimes develop and consider the hypothesis that, if we were to change our narratives to either more-closely or less-closely resemble the narrative described, we might lead better lives, ourselves. Works of literature can sometimes give us a sense of how we might see the world differently, and of the kinds of narratives we might construct for ourselves if we did.

Fictional literature often presents narratives involving characters who focus on different properties of objects and situations from those we typically focus on, and who have different skills of pattern recognition from ours. Literature often enables readers to take on the perspectives and phenomenology of these characters (Farrell 2007: 256; cf. Elgin 2007: 51). Elgin proposes, for example, that if we took on the "morally impeccable point of view" of the character Fanny Price in Austen's *Mansfield Park*, and learned to "focus on the sorts of factors" her view discloses, we might be able to "increase our own moral sensitivity" in everyday life (2007: 52). This essay adds that, having changed our perspective, we might also consider constructing new, future-oriented narratives for ourselves that we may not have considered, otherwise. Elgin writes that even if a character's perspective is not accurate, it can still be rewarding to adopt it. For instance, she notes that, by taking on the view of Holden Caufield in Salinger's *The Catcher in the Rye*, we might come to recognize things as phony that we might otherwise fail to rightly see as hypocritical (Elgin 2007: 52).

Without the help of literature, it might often be difficult to fully understand perspectives that are very different from our own. According to Karen Simecek, a perspective is: "a set of implicit beliefs, commitments, and values" that determines "not only what information becomes the focus of our thinking; i.e. *what* we will bring to the fore, but *how* we will organize that information in bringing those features to the fore" (2015: 501). Elisabeth Camp notes that a perspective or-

ganizes our thoughts, making some features of the world "stick out in our minds" while others do not (2009: 111); changing perspective can change what we focus on, how we categorize what we perceive, and the patterns we notice.

In real life, we engage in what some theorists call mindreading—attributing mental states to people we encounter and coming to believe that they have specific beliefs and feelings (cf. Goldman 2005; Gallagher 2011: 21–22; 2012: 356 and 359; 2014: 601). But, while it might sometimes be possible to have a sense of some of people's mental states, it is often difficult to fully grasp their perspectives. It is difficult to understand just how people think of their circumstances, on what details and information they focus, or the kinds of patterns they notice. Suppose we try to imaginatively put ourselves in another person's shoes, aiming to set aside our own mental states and substitute those of the target (cf. Goldman 2005: 86; cf. 79–81). Then suppose we ask ourselves how we would think of the circumstances the target faces, were we in his/her shoes. How reliably can we set aside our own mental states and adopt someone else's? How often would we mistakenly suppose that the target would perceive and approach his/her circumstances much like we would approach them if we faced those circumstances, ourselves, when in fact the target, whose perspective is radically different from ours, would approach circumstances much differently (cf. Gallagher 2012: 363–364 and 370; 2014: 602–603)? As Shaun Gallagher remarks, if we rely on our own experiences and imagination, it would be difficult to take into account other people's vast diversity of beliefs, perspectives, and experiences (ibid.).

Yet literary narratives can give us an expanded understanding of diversity, introducing characters who are unlike anyone we have met, events we have never experienced, and locations we have never visited (Nussbaum 1990: 47–48). Literary narratives provide a sense of how people very different from us might behave in a variety of circumstances, what reasons they might have for their actions, and how they might feel (cf. Gallagher 2011: 31–33; 2012: 371). Gallagher maintains that an education in narratives can provide us with background knowledge about people very different from us, on the basis of which we can better understand the perspectives of people in real life and better predict how they might think, feel and act (2011: 31–34; 2012: 371–372; 2014: 605–606). We can also take works of fictional literature as offering hypotheses about which actions would be considered socially acceptable in a variety of circumstances, how we, ourselves, could choose to act in those circumstances, and about what would constitute good reasons for different actions (cf. Gallagher 2012: 371; 2014: 605–606). When we better understand what actions are possible and acceptable in a variety of circumstances, we gain a better sense of possible future-oriented personal narratives we could develop for ourselves.

As Simecek notes, literary narratives often draw attention to specific details of the characters they describe, suppress other details, categorize

and compare characters, and encourage readers to adopt specific perspectives of characters (2015: 502–504 and 508). The details on which works of literature focus, and the categorizations they make, might be very different from the details and categorizations on which readers would focus if they encountered people like the works' characters in real life (Camp 2009: 117). Camp indicates that once readers gain a sense of the different details on which they could focus, categorizations they can make, and perspectives they could take, they might change what they focus on in the real world and how they interpret what they encounter (Camp 2009: 117). Even if we cannot adequately understand or adopt the perspectives of some real people we meet, we might still be able to at least partially understand and adopt the perspectives offered by some fictional literary narratives. These new perspectives, in turn, might enable us to develop new personal narratives for ourselves.

Kivy notes that if a deep and complex fictional character that we care about and understand holds an opinion, then perhaps we will take that opinion more seriously (Kivy 2006: 113). This essay adds that if fictional characters like that have narratives very different from our own, we might be more inclined to consider the hypothesis that we should create future-oriented narratives more like theirs. We might also be more likely to see this hypothesis as living. Likewise, when we appreciate these characters' distinctive perspectives, we might better understand how we would need to see the world in order to live according to future-oriented narratives like theirs. Also, as Novitz indicates, when we read about these characters applying practical or intellectual strategies to achieve their goals, and vividly imagine them doing so, we might reflect on the hypothesis that similar strategies might be able to help us reach similar goals—goals that we might otherwise have thought were impossible for us (1984: 49 and 57–60).

It is beyond the scope of this essay either to carefully explore when readers would be justified to consider adopting new personal narratives in response to reflecting on hypotheses raised by works of fiction, or to explore how someone who reads numerous works of fiction might reasonably choose which of their hypotheses to test out in real life. Imagine a reader who is wondering which hypotheses to test out in real life, and is confident that testing out any of them would not be immoral.[6] Perhaps future research could consider whether it is justifiable for a reader like this to consider choosing to test out just those hypotheses the reader feels are particularly emotionally compelling. How often do even scientists who apply good research methodology rely on emotions and gut-feelings when determining which hypotheses to test?[7] Separately, it is also beyond the scope of this essay to provide an account

[6] For content relevant to the problem of determining which hypotheses would be moral, see: Nelson 2001: 41–45 and 51–52.

[7] Consider Alexander Klein's reflections on William James's approach to the role of one's passions when selecting which hypotheses to consider (2015: 84–92).

of exactly what a work of literature and reader would need to be like, for the work to prompt the reader to consider adopting a new personal narrative. Yet, in the next section, this essay will consider an example of a work of literature—*10:04*—that can prompt readers to consider the hypothesis that they should change their personal narratives.

3. *A Case Study: Learning from Lerner's 10:04*

Ben Lerner's novel, *10:04,* is narrated by a character who, like Lerner himself, is an author living in Brooklyn and who describes his pressures and life experiences during the time in which he hopes to write a new novel. The narrator describes a number of personal challenges he faces: he has just heard from his doctor that he has a potentially life-threatening condition, and his best friend has asked if she could use his sperm for artificial insemination. The narrator also focuses on his day-to-day social interactions, and describes trivial concerns of day-to-day life that can nonetheless induce stress. For example, the narrator at one point chaperones a boy on a visit to a museum; Lerner provides a four-page description of a five-minute scene in which the narrator needs to find a restroom, panics about leaving the child on his own, speaks with the boy about where to wait for him, and then panics when the boy is not exactly where he had promised to remain (2014a: 149–152).

David James mentions Lerner as an example of a "localist" writer (James 2017: 133)—the kind of author who focuses substantially on "quotidian circumstances" and the "fiction of local life" (James 2015: 47). Yet James is careful to note that localist fictions, like *10:04,* can nonetheless engage with "worldly concerns" (2015: 49), and can explore the tensions and relations between local and global issues. As much as *10:04* focuses on the day-to-day, localized life of its narrator, it also focuses on global concerns, like global warming, and reflects on great expanses of time.

Ben De Bruyn describes *10:04* as an "example of a contemporary climate-change novel" because, while it provides "a remarkably realist attention to everyday life," it frequently reflects "on our increasingly unrecognizable planet" and its narrator's anxieties about global environmental crises (2017: 951; cf. Gibbons 2019a; 2020: 1). The narrator, who is worried about global warming, refers to the weather's "unseasonable warmth" repeatedly throughout the novel (2014a: 3, 32, 63, 66, 107, 164, 206, 213, 221, and 231; cf. Gibbons 2019a; Gibbons 2020: 10). Likewise, the narrator often imagines New York City either as "underwater" or as "sinking" (2014a: 4, 40, and 153; Gibbons 2020: 10). The narrator has a conversation with a boy—the same he chaperoned at the museum—who has had a dream that "the buildings all freeze up after global warming makes an ice age" (2014a: 13), and the narrator empathizes with the boy whom he says, "like me, tended to figure the global apocalyptically" (2014a: 14; cf. Gibbons 2019a: 2020: 10).

The novel's narrative is bookended by hurricanes Irene and Sandy, which, as Alison Gibbons notes, grounds the novel in historical time (2019a; cf. 2020: 5–6; De Bruyn 2017: 953). As De Bruyn indicates, the novel's narrative also frequently refers to "the inhuman timeframes of paleontology and astronomy" (De Bruyn 2017: 953). For example, the narrator reflects at substantial length upon his recognition that the brontosaurus might never have existed and that Pluto might not be a planet—recognitions that challenge his "remembered sense of both galactic space and geological time" (Lerner 2014a: 11; De Bruyn 2017: 953). Likewise, while the narrator reflects at length on the specific events of his day-to-day life, he sometimes considers the broader contexts and global networks that have made events in his life possible. When he eats octopus at a restaurant, for instance, he briefly muses about "the rhythm of artisanal Portuguese octopus fisheries … the mercury and radiation levels of the sashimi and the chests of the beautiful people in the restaurant—coordinated, or so it appeared, by money" (Lerner 2014a: 156; De Bruyn 2017: 957).

The narrator in *10:04* might serve as an example to readers of how they might choose to lead their everyday lives, all the while aware of the environmental crises affecting the planet. Each of us is going to continue to lead a full life, complete with everyday social interactions and personal challenges. How might we best think about how to lead our local lives, even as we are aware of global crises? Lerner implies that he had questions like this—if not this exact question—in mind, himself. In an interview, he remarks:

> I can see why, if I, the historical person, choose to write a book that's set in Brooklyn that talks about book advances and eating Bluefin tuna or whatever, that it's automatically in the category of the self-absorbed … The book wants to acknowledge all of that as an attempt to see what spaces for healing can exist, as opposed to the model of fiction that's like "The way I deal with the political is that I pretend to have access to the mind of a nine-year old boy in Sudan" (Lerner 2015).[8]

After reading *10:04,* we might develop for ourselves, and reflect on, the hypothesis that it could be beneficial for us to develop future-oriented narratives that take into account and value both local, everyday concerns, and considerations about global crises and the future of the planet. While *10:04* does not explicitly provide this hypothesis, some readers might develop it and reflect on it. Readers can compare themselves with the narrator of *10:04,* and consider how they, like the narrator, might develop a narrative that balances local and global concerns.

10:04 provides an example of what such a narrative could resemble, giving us the perspective of a narrator who focuses on information and recognizes patterns both in his local circumstances and in global issues. Gibbons argues that fiction has the ability to ask readers to "think historically—to place ourselves within a larger narrative of geo-

[8] James provides this quotation when discussing how Lerner's localist fiction engages with global concerns (2017: 134–135).

logical time and human evolution" (Gibbons 2019b: 293). This essay contends that, after reading *10:04*, some readers might reflect that it is sometimes beneficial to think historically, but that it is also valuable to focus on and care about everyday concerns. Readers might hypothesize about the value of balancing local and global concerns in their personal narratives, and they might choose to adopt a perspective and personal narrative more similar to that of the narrator of *10:04*.

Readers might think that the narrator of *10:04* does not do enough to alleviate the world's environmental concerns. Yet the narrator's insufficiency, itself, might prompt readers to actively consider what they could do, themselves, to more successfully address those concerns. If the narrator had no shortcomings, some readers might never reflect on what they could do in comparison. Suppose that, having read *10:04,* readers consider the hypothesis that it would benefit them to live more like the narrator than they currently do. Even if readers reject this hypothesis, because the narrator does not do enough to address environmental concerns, the readers' reflections might still have valuably influenced how they see their place in the world.[9]

As Iris Vidmar notes, a hypothesis that we extract from a work of fiction is not valuable only if we determine it true and if it results in us acquiring new knowledge; rather, part of its value is determined by how it impacts the ways we view ourselves and our experiences (2013: 190). Consider novels that present flawed characters and hypotheses that readers reject. In some cases, those novels might lead readers to engage in reflection long after they have finished reading—reflections they might not have had if they had approved of the characters and had agreed with the novels' hypotheses. Such reflections might valuably shape how the readers view themselves and their experiences, in comparison with the novels' characters.

4. *Authoring a fiction, authoring a life*

10:04 is a particularly good example of a novel that highlights the impact that fictions can have on their readers. Lerner notes that *10:04* is meant to reflect in part on "how fiction functions in our real lives" (Lerner 2014B). He adds: "My concern is how we live fictions, how fictions have real effects, become facts in that sense, and how our experience of the world changes depending on its arrangement into one narrative or another" (2014B; cf. Gibbons 2018: 86 and 94). While Lerner might not have had in mind anything quite like what this essay suggests, his thoughts in this quotation are nonetheless relevant.

What makes *10:04* a good example, when exploring the ways in which fictions can affect how we lead our factual lives, is that *10:04* belongs to a literary genre known as "autofiction" (van den Akker, Gib-

[9] Iris Vidmar writes that often the hypotheses "we extract from works … will not provide definite answers, but will certainly influence the way we think about … our place in the world" (2013: 190).

bons and Vermeulen 2019: 48–49; Gibbons 2018: 75–76 and 85). As a work of autofiction, *10:04* mixes elements of autobiography with fiction, and focuses on the relation between them.

As Robin van den Akker, Alison Gibbons, and Timotheus Vermeulen note, the narrator of *10:04* shares the same name as the author (2019: 49), and the novel is a blend of fiction and autobiography, drawing directly from the author's real life and work. For example, according to the novel, the narrator has written a short-story which is included, in full, in *10:04*—a story that is, in fact, a republication of a story that the real-world Ben Lerner had published in *The New Yorker* titled, "The Golden Vanity" (Gibbons 2018: 85). Like the author, the narrator also undertakes a residency in Marfa, Texas, during which the narrator writes poetry that is identical to that of the author (Lerner 2014a: 172–176). As a result, *10:04* sometimes seems to encourage its audience to see the novel's narrator as the same person as its real-world author (van den Akker, Gibbons and Vermeulen 2019: 49).

That said, in interviews the author is clear about the extent to which *10:04* is a work of fiction. For example, he writes that a children's book which the narrator co-authored with a child he tutors "is based on a book with a great kid I tutored, but the 'Roberto' character doesn't resemble him very much" (Lerner 2014b). Additionally, a central relationship in *10:04* "is the relationship between the narrator and Alex. Alex isn't based on a real person, at least not a single real person" (Lerner 2014b). van den Akker, Gibbons, and Vermeulen write: "Ben in *10:04* flickers between being an inhabitant of a constructed novel and that of a seemingly depthy real world" (2019: 49; cf. Gibbons 2018: 85–86).

10:04 demonstrates our ability to create new narratives for ourselves; Lerner has written a book according to which he, as the narrator, leads a very different life from the one that he, as the author, has lived. The novel is an example of a constructed narrative, invented by a real person which is—at least in a loose sense—about that real person, too. If the narrator did not seem like "an inhabitant of a constructed novel" (van den Akker, Gibbons and Vermeulen 2019: 49), then *10:04* would not be an example of a *constructed* narrative, and if the narrator did not seem like an inhabitant of a "depthy real world" (van den Akker, Gibbons and Vermeulen 2019: 49), then *10:04* would not be an example of a new narrative that is invented by a *real* person that is also about that real person.

The novel does not provide a narrative that Lerner could actually choose to act out in real life. Yet there are clear parallels between authoring an autofictional narrative and constructing a future-oriented personal narrative for oneself. In both cases, the individual constructing the narrative is both "a figure within the story and the person writing it" (Lerner 2012). In both, an individual "invents" a narrative that might deviate in part from true descriptions of the real world.[10]

[10] Velleman discusses the role of individuals as inventors of narratives and the relation between personal narratives and fictions (2005: 71–72).

Consider an individual who, at the end of a day of work, says, "I am going home now," and constructs a mental narrative in which he does, fully intending to follow through. But suppose that, either as a result of weakness of will or outside circumstances, he fails to act out the narrative he developed. Like a work of autofiction, the protagonist of the narrative would in many ways resemble (or *be*) the author of the narrative, but, like a work of autofiction, many of the narrative's details would differ from reality.[11] When individuals construct future-oriented narratives according to which they intend to live, they could not know for sure to what extent their narratives resemble autofiction; those individuals could not know if they will successfully live as they intend. Lerner might not have had this sort of comparison between narratives and autofiction in mind, but, in an interview about his story, "The Golden Vanity," which later became part of *10:04*, he does remark:

> "[T]he author" begins to feel both like a figure within the story and the person writing it—we're both trying to figure out how he can continue … I'm talking about something intensely lived: How each of us is constantly striving to reorganize mere chronology into some meaningful pattern, to narrate our pasts in a way that makes a future thinkable. The part of the cliché "you're the author of your own life" that I agree with is its implication that our identities are fictions. (Lerner 2012)

Reflecting on *10:04*—and on Lerner's ability to construct a fictional narrative about himself—might encourage readers to consider their own ability to construct new personal narratives, some of which they might aim to put into action in real life. In fact, *10:04* frequently reminds us of our ability to invent narratives for ourselves (van den Akker, Gibbons and Vermeulen 2019: 51), as when the narrator realizes that the story he has told himself, about how he became a poet, does not cohere with his real-life experiences and is based in part on mis-remembered facts (2014a: 110–116; cf. Bilmes 2018: 3). *10:04* also frequently indicates that the future is open to us; there are numerous possibilities to pursue and narratives we can create (cf. Gibbons 2020: 2 and 8–11). For example, the narrator reflects on the possibility of writing a novel in which he will "project myself into several futures simultaneously" (2014a: 4; cf. 194; Bilmes 2018: 6–10; Gibbons 2019a; 2020: 11).

Earlier, this essay suggested that people should view their personal narratives from three different perspectives: that of author, character, and critic. *10:04* reflects on a figure—Ben Lerner—who is all three; he is the author, who sometimes, like a critic, reflects on his own authorial choices, and he is also a character about whom he writes—a character who himself is also an author. *10:04* might also encourage its audience to reflect on how to juggle these three perspectives when approaching their own life narratives.

Insofar as *10:04* encourages audience members to consider the ways

[11] While Velleman does not discuss the notion of autofiction, he does likewise considers this example in relation to the notions of personal narratives, fact, and fiction (2005: 71–72).

in which their local, personal lives relate to global issues, like global warming and environmental crises, it might encourage readers to consider revising their personal narratives to take these crises into account. Perhaps, as Gibbons suggests is possible for fiction, *10:04* can express: "a refusal to accept the current state of the world, asking readers instead to think critically and defiantly about the ways in which world events are connected and how their own involvement figures in such a world" (Gibbons 2015: 41; cf. van den Akker, Gibbons, and Vermeulen 2019: 52). Perhaps, in order to most effectively address global issues, like the world's environmental crisis, we will also need to revise our personal narratives like this (Gare 1995: Chapter 5; 2001; cf. Kirkpatrick 2014: 60 and 64–65; Corsa 2018: 247–249).

References

Bilmes, L. 2018. "'An Actual Present Alive with Multiple Futures': Narrative, Memory and Time in Ben Lerner's *10:04*." *Textual Practice,* advance-publication: 1–22.

Bruner, J. 1991. "The Narrative Construction of Reality." *Critical Inquiry* 18: 1–21.

Bruner, B. and Kalmar, D.A. 1998. "Narrative and Metanarrative in the Construction of Self." In M. Ferrari and R.J. Sternberg (eds.). *Self-Awareness: Its Nature and Development*. New York and London: The Guilford Press: 308–331.

Camp, E. 2009. "Two Varieties of Literary Imagination: Metaphor, Fiction, and Thought Experiments." *Midwest Studies in Philosophy* 33: 107–130.

Carr, D. 1986. *Time, Narrative, and History*. Bloomington/Indianapolis: Indiana University Press.

Corsa, A. J. 2018. "Grand Narratives, Metamodernism, and Global Ethics." *Cosmos and History: The Journal of Natural and Social Philosophy* 14 (3): 241–272.

Davies, D. 2007. "Thought Experiments and Fictional Narratives." *Croatian Journal of Philosophy* 7 (19): 29–45.

De Bruyn, B. 2017. "Realism 4°. Objects, Weather and Infrastructure in Ben Lerner's *10:04*." *Textual Practice* 31 (5): 951–971.

Dianda, A. 2018. "William James and the 'Willfulness' of Belief." *European Journal of Philosophy* 26: 647–662.

Elgin, C. Z. 2002. "Art in the Advancement of Understanding." *American Philosophical Quarterly* 39 (1): 1–12.

Elgin, C. Z. 2007. "The Laboratory of the Mind." In J. Gibson, W. Huemer, and Luca Pocci (eds.). *A Sense of the World: Essays on Fiction, Narrative, and Knowledge.* New York and London: Routledge: 43–54.

Farrell, F. B. 2007. "'The way light at the edge of a beach in autumn is learned': Literature as Learning." In J. Gibson, W. Huemer, and Luca Pocci (eds.). *A Sense of the World: Essays on Fiction, Narrative, and Knowledge.* New York and London: Routledge: 246–265.

Gallagher, S. 2011. "Narrative Competence and the Massive Hermeneutical Background." In P. Fairfield (ed.). *Education, Dialogue and Hermeneutics*. London and New York: Continuum: 21–38.

Gallagher, S. 2012. "Empathy, Simulation, and Narrative." *Science in Context* 25: 355–381.

Gallagher, S. 2014. "An Education in Narratives." *Educational Philosophy and Theory* 46 (6): 600–609.

Gare, A. 1995. *Postmodernism and the Environmental Crisis*. London and New York: Routledge.

Gare, A. 2001. "Narratives and the Ethics and Politics of Environmentalism: The Transformative Power of Stories." *Theory and Science* 2 (1): http://theoryandscience.icaap.org/content/vol002.001/04gare.html

Gare, A. 2002. "Narratives and Culture: The Role of Stories in Self-Creation." *Telos* 122: 80–100.

Gibbons, A. 2015. "'Take that you intellectuals!' and 'kaPOW!': Adam Thirwell and the Metamodernist Future of Style." *Studia Neophilologica* 87: 29–43.

Gibbons, A. 2018. "Autonarration, I, and Odd Address in Ben Lerner's Autofictioanl Novel *10:04*." In A. Gibbons and A. Macrae (eds). *Pronouns in Literature: Positions and Perspectives in Language*. London: Palgrave Macmillan: 75–96.

Gibbons, A. 2019A. "Metamodernist Futures or 'a Future that Had Never Arrived': Climate Change and Temporality in *10:04."* Paper presented at *Metamodernism*, Radboud University, Nijmegen, Netherlands.

Gibbons, A. 2019B. "Entropology and the End of Nature in Lance Olsen's *Theories of Forgetting*." *Textual Practice* 33 (2): 280–299.

Gibbons, A. 2020. "Metamodernism, the Anthropocene, and the Resurgence of Historicity: Ben Lerner's *10:04* and 'The Utopian Glimmer of Fiction.'" *Critique: Studies in Contemporary Fiction*, advance publication: 1–15.

Goldie, P. 2012. *The Mess Inside: Narrative, Emotion, and the Mind*. Oxford: Oxford University Press.

Goldman, A.I. 2005. "Imitation, Mind Reading, and Simulation." In S. Hurley and N. Chater (eds.). *Perspectives on Imitation: From Neuroscience to Social Science*. Vol 2. Cambridge: The MIT Press: 79–93.

Greenwald, A.G., Carnot, C.G., Beach, R., and Young, B. 1987. "Increasing Voting Behavior by Asking People if they Expect to Vote." *Journal of Applied Psychology* 72 (2): 315–318.

Hutto, D.D. 2016. "Narrative Self-Shaping: A Modest Proposal." *Phenomenology and the Cognitive Sciences* 15: 21–41.

James, D. 2015. "Worlded Localisms: Cosmopolitics Writ Small." In L. Platt and S. Upstone (eds.). *Postmodern Literature and Race*. Cambridge: Cambridge University Press: 47–61.

James, D. 2017. "Transnational Postmodern and Contemporary Literature." In Y. Goyal (ed). *The Cambridge Companion to Transnational American Literature*. Cambridge: Cambridge University Press: 122–140.

James, W. 1889. "The Will to Believe." In *The Will to Believe and Other Essays in Popular Philosophy*. New York, London, and Bombay: Longman Green and Co: 1–31.

Kirkpatrick, A. 2014. "Modernity, Post-Modernity and Proto-Historicism: Reorienting Humanity through a New Sense of Narrative Emplotment." *Cosmos and History: The Journal of Natural and Social Philosophy* 10 (2): 22–77.

Kivy, P. 2006. *The Performance of Reading: An Essay in the Philosophy of Literature*. Malden: Blackwell Publishing.

Klein, A. 2015. "Science, Religion, and 'The Will to Believe.'" *HOPOS: The Journal of International Society for the History of Philosophy of Science* 5 (1): 72–117.

Lamarque, P. 2007. "On the Distance between Literary Narratives and Real-Life Narratives." In D. D. Hutto (ed.). *Narrative and Understanding Persons*. Cambridge: Cambridge University Press: 117–132.

Lerner, B. 2012, June 10. "This Week in Fiction: Ben Lerner." Interview by C. Leyshon. *The New Yorker:* https://www.newyorker.com/books/page-turner/this-week-in-fiction-ben-lerner

Lerner, B. 2014a. *10:04: A Novel*. New York: Faber and Faber.

Lerner, B. 2014b, September 1. "An Interview with Ben Lerner." Interview by T. Lin. *The Believer*, 110: https://believermag.com/an-interview-with-ben-lerner/

Lerner, B. 2015, January 3. "People say, 'Oh, Here's Another Brooklyn Novel by a Guy with Glasses." Interview by E. Witt. *The Guardian:* https://www.theguardian.com/books/2015/jan/03/ben-lerner-1004-novel-books-interview

MacIntyre, A. 2008. *After Virtue: A Study in Moral Theory*. Notre Dame, IN: University of Notre Dame Press.

Nelson, H. L. 2001. *Damaged Identities, Narrative Repair*. Ithaca, NY and London: Cornell University Press.

Novitz, D. 1984. "Fiction and the Growth of Knowledge." In J. Margolis (ed.). *The Worlds of Art and the World*. Amsterdam: Rodopi: 47–68.

Nussbaum, M. C. 1990. *Love's Knowledge: Essays on Philosophy and Literature*. Oxford: Oxford University Press.

Robinson, G. C. 2014. "You Live and Learn: Narrative in Ethical Enquiry with Children." *Childhood and Philosophy* 10 (20): 305–330.

Simecek, K. 2015. "Beyond Narrative: Poetry, Emotion and the Perspectival View." *British Journal of Aesthetics* 55 (4): 497–513.

Schechtman, M. 2011. "The Narrative Self." In S. Gallagher (ed.). *The Oxford Handbook of the Self*. Oxford: Oxford University Press: 394–416.

Tollefsen, D. and Gallagher, S. 2017. "We-Narratives and the Stability and Depth of Shared Agency." *Philosophy of the Social Sciences* 47 (2): 95–110.

van den Akker, R., Gibbons, A., and Vermeulen, T. 2019. "Metamodernism: Period, Structure of Feeling, and Cultural Logic—A Case Study of Contemporary Autofiction." In D. Rudrum, R. Askin, and F. Beckman (eds.). *New Directions in Philosophy and Literature*. Edinburgh: Edinburgh University Press: 41–54.

Velleman, J. D. 2005. "The Self as Narrator." In J.P. Christman and J. Anderson (eds.). *Autonomy and the Challenges to Liberalism: New Essays*. Cambridge: Cambridge University Press: 56–76.

Vice, S. 2003. "Literature and the Narrative Self." *Philosophy* 78: 93–108.

Vidmar, I. 2013. "Though Experiments, Hypotheses, and Cognitive Dimension of Literary Fiction." *Synthesis Philosophica* 55–56 (1–2): 177–193.

Croatian Journal of Philosophy
Vol. XXI, No. 61, 2021
https://doi.org/10.52685/cjp.21.1.7
Received: July 18, 2020
Accepted: November 12, 2020

Doing Things with Words: The Transformative Force of Poetry

PHILIP MILLS
University of Lausanne, Lausanne, Switzerland

Against the apparent casting away of poetry from contemporary philosophy of language and aesthetics which has left poetry forceless, I argue that poetry has a linguistic, philosophical, and even political force. Against the idea that literature (as novel) can teach us facts about the world, I argue that the force of literature (as poetry) resides in its capacity to change our ways of seeing. First, I contest views which consider poetry forceless by discussing Austin's and Sartre's views. Second, I explore the concept of force in the realm of art—focusing on Nietzsche's philosophy and Menke's Kraft der Kunst*—and the relations between linguistic, artistic, and political forces. Third, I consider how the transformative force of poetry can be considered political by turning to Kristeva's* Revolution in Poetic Language *and Meschonnic's conception of poetry according to which the poem does something to language and the subject. To illustrate this transformative force of poetry, I analyse Caroline Zekri's poem 'Un pur rapport grammatical'. I therefore think of poetry not only as doing something with language, but also as doing something to language. To rephrase Austin's famous title, and thus reverse his evaluation of poetry, poetry might reveal us not only* How to Do Things with Words*, but how to do things to words and, through this doing, how to transform and affect the world.*

Keywords: Philosophy of poetry; ordinary language philosophy; force.

> *Je suis donc obligé de dire que le poème fait quelque chose. Il fait quelque chose au langage, et à la poésie. Il fait quelque chose au sujet. Au sujet qui le compose, au sujet qui le lit.* (Meschonnic 2001: 43)[1]

Despite its attempt to systematically analyse the specificities of each artform, contemporary aesthetics seems to have surprisingly left poetry aside. From being the paragon of the arts in 18th and 19th century philosophy—as in Baumgarten, Kant, or Hegel for instance—poetry in the contemporary world seems to have lost most, if not all, of its philosophical force. Even Plato, who is famous for being rather unkind to poetry, nevertheless admits that poetry has a specific force, one he is afraid of, and his unkindness reveals his fear of poetry rather than an indifference towards it. In contemporary aesthetics, as John Gibson argues, 'until very recently one could fairly say that poetry is the last great unexplored frontier in contemporary analytic aesthetics, an ancient and central art we have managed to overlook more or less entirely' (Gibson 2015: 1). Even though one might argue that the situation has changed since then and that Gibson overlooks some traditions in philosophy's dealing with poetry, it seems that poetry has lately received less attention than film or the novel for instance. How can one explain such a change of attitude towards poetry? One of the main reasons for this shift can be found in one of the grounds of analytic philosophy: the linguistic turn'. If, following this turn, philosophy is a matter of language and solving philosophical problems becomes a matter of solving linguistic ones, poetry seems to be of no help at all, quite to the contrary, to philosophy of language either as 'ideal language philosophy' or as 'ordinary language philosophy', the two types of philosophy of language Rorty considers in editing *The Linguistic Turn* (Rorty 1967: 15). If poetry is a problem for the former, as it presents a form of language which cannot be translated into formal logic and therefore not be given any truth-value, the latter also shows no interest in it, as Austin suggests that performative utterances in a poem, are '*in a peculiar way* hollow or void' (Austin 1975: 22). Failure for philosophy of language to give a substantial account of the language of poetry might have contaminated the realm of aesthetics and incited philosophers to look at artforms other than poetry, more easily approachable with these new philosophical tools.[2] The great interest in literature (here understood as the novel)

[1] My translation: 'I am therefore forced to say that the poem does something. It does something to language and to poetry. It does something to the subject. To the subject that composes it, to the subject that reads it.'

[2] It must be said that most philosophies of language have attempted to give an account of poetic phenomena but have often considered poetic language from the sole perspective of metaphor. Metaphor has been of importance throughout the whole history of philosophy of language, but inasmuch as metaphor is only one poetic phenomenon among many others, these accounts of poetic language do not exhaust the possibilities used in poetry. It is worth noting that metaphors pervade our everyday speech (Lakoff and Johnson 2003) and can therefore not really be considered a definitory feature of poetic language. This idea that metaphor is central

and the problem of truth in fiction can be seen as a consequence of the 'linguistic turn': philosophers have started looking into aesthetic problems for which philosophy of language could be of use, rather than artforms which are problematic to philosophy of language.[3] Inasmuch as Austin deprives poetry from any performative force, contemporary aesthetics strips poetry from its philosophical force.

If this account gives a schematic picture of the place of poetry in analytic aesthetics, one might think poetry fares better on the other side of the so-called 'analytic-continental divide.' At first glance, continental philosophers seem to pursue the 19th century praise of poetry.[4] Heidegger for instance considers poetry almost on par with philosophy, and the most philosophical of all artforms. If one looks further, however, a shift in attitude similar to that of analytic aesthetics seems to occur in continental philosophy. Although it is not a mark of indifference towards poetry, Sartre's theory of literature seems to operate a similar shift from poetry to literature (and one can understand here, as with analytic philosophy, 'the novel'). Indeed, Sartre defines literature in terms of political commitment and denies any commitment to poetry.[5] Very schematically: if the force of literature is proportional to its political commitment, poetry has no force as it is denied such commitment. This does not mean that Sartre denies any greatness to poetry but ascribes it one which might be of another kind, one which is certainly not of help to any concern in the actual world. If poetry still has a force in continental aesthetics, it is not a political one, not a force of influence in

to poetry (and it is in some cases) shows a focus on 19th century poetry rather than on contemporary forms. If a philosophy of poetry is to give an account of current poetic practices as well, metaphor might not be of great help.

[3] Contemporary philosophy of literature indeed shows little interest in poetry, perhaps because the notion of fiction which is central to studies in literature does not apply to poetry, or not in a straightforward way. One could even say that poetry undermines the fiction/nonfiction divide. Contemporary autofictions also work towards this undermining, but as for poetry, philosophy of literature seems to be exclusively focused on the 19th century novel, as Peter Lamarque argues: 'Philosophy of Literature has virtually become Philosophy of the Novel' (Lamarque 2017: 109).

[4] It seems indeed that all continental philosophers bring their attention to poetry in some way or another. However, this (sometimes) central place they give to poetry further attempts to isolate it from the ordinary world, to leave it in its ivory tower.

[5] Maximilian de Gaynesford offers an interesting reading of Austin in relation to the notion of commitment (although not necessarily political commitment), which opens a potential link between Austin and Sartre. According to him, either poetry is capable of commitment and is thus serious, or it is not and can therefore not claim to seriousness: 'For if we insist that poetry is "serious", Austin can still rescue his approach to speech acts; he must simply accept that commitment-apt utterances in poetry may make commitments. And it is surely possible to contemplate ways of integrating poetry and philosophy while acknowledging that poetry is, indeed, serious. (The surprise is that we might have been able to do so without acknowledging this.) We would, however, have to renounce the attempt to exempt poetry from forms of commitment' (de Gaynesford 2011: 49). We then must choose whether we would rather commit poetry to seriousness and hence abandon the idea of poetic license or abandon commitment in favour of poetic license.

the everyday world, but rather a force of distancing and isolating itself from the commonplace.

One of the possible reasons for this shift is an inversion of value between literature and poetry. Whereas poetry was literature (or the highest literary form) for 18th and 19th century philosophers (and in this sense Heidegger inherits from this background and pursues a romantic tradition), the 20th century marks the rise of the novel in philosophical concerns. When one thinks of literature nowadays, the first thing to come to mind is probably more often a novel than a poem. In that sense, philosophy of poetry would be considered a subcategory of philosophy of literature rather than the opposite.[6] However, even if there were such a shift, it would not explain the disdain towards poetry and why philosophers have stripped it from its force. In my paper, I therefore aim at reinstating the force of poetry by showing that it has a linguistic, philosophical, and even political force (and this as much as the novel) because poetry, I will argue, has a revolutionary dimension. Against the idea that literature (as novel) can teach us facts about the world, I argue that literature (as poetry) teaches us a different way of seeing the world and that its force resides precisely in its capacity to bring us (or force us, perhaps) to see things differently. As Wittgenstein puts it: 'The work of art compels us to see it in the right perspective' (Wittgenstein 1998: 7) More than seeing the work of art itself in the right perspective, it compels us to see the world in the right perspective, in a perspective which makes sense.

To explore the force(s) of poetry, my paper is divided in three parts. First, I contest views which consider poetry forceless, be it linguistically or politically, by discussing Austin's and Sartre's views. Second, I explore the concept of force in the realm of art—focusing on Nietzsche's philosophy and Menke's *Kraft der Kunst*—and the relations between linguistic, artistic, and political forces. Third, I explore how the transformative force of poetry can be considered political by turning to Kristeva's *Revolution in Poetic Language* and Meschonnic's conception of poetry according to which the poem does something to language and to the subject. To illustrate this transformative—and even political—force of poetry, I offer an analysis of Caroline Zekri's poem 'Un pur rapport grammatical'. Poetry is revolutionary in the sense that it brings to see language and the world anew. Poetry is not only doing something with

[6] Peter Lamarque for instance considers philosophy of poetry to be a sub-branch of philosophy of literature: 'But just when aestheticians are getting used to another recent sub-division, The Philosophy of Literature, here comes a sub-branch of that, focused on poetry' (Lamarque 2017: 109). We could argue against Lamarque that philosophy of literature is not so recent (and perhaps one of the oldest forms of philosophy of art if we think of Aristotle) and that philosophy of poetry, historically speaking, comes first. However, I agree with Lamarque to a certain extent in considering philosophy of poetry distinct from philosophy of literature. My only contention is that philosophy of poetry must be understood as radically different from philosophy of literature rather than as a sub-branch of it.

language, but also doing something to language. To rephrase Austin's famous title, and thus reverse his evaluation of poetry, poetry might not reveal us *How to Do Things with Words*, but how to do things *to* words. The force of poetry is not primarily political, but it becomes political insofar as its force modifies language and, through this modification of language, our ways of being in the world.

1. *The forcelessness of poetry: Austin and Sartre*

This first section explores two ways in which poetry is considered forceless, respectively from a linguistic and a political perspective. As outlined in the introduction, Austin considers poetic statements to be without performative force and Sartre poems to lack political commitment. Although the authors are quite distant from one another, these two considerations are not unrelated as Sartre's argument against the political force of poetry relies on depriving poetry from any linguistic force.

In *How to Do Things with Words*, Austin distinguishes between constative utterances, statements of fact, and performative utterances 'in which *by* saying or *in* saying something we are doing something' (Austin 1975: 12). In other words, and following Austin's example, saying 'I promise' is not a statement of fact but a performance of the act of promising, as much as saying 'I do' in a marriage ceremony is a performance of the act of marrying. 'Here we should say that in saying these words we are *doing* something—namely, marrying, rather than *reporting* something, namely *that* we are marrying' (Austin 1975: 13). There is a distinction between speech acts that describe or report a fact and speech acts that perform an action, although both kinds of utterances might look alike. Performative utterances might grammatically look like statements, but they are different from them in that they cannot be said to be true or false; they have no truth-value, but they have a certain force.

At first glance, replacing the notion of truth-value by that of force might open the possibility to account for uses of language which escape the game of truth-value, such as poetic and literary uses. Indeed, in the framework of conceptions of language based on truth-value, poetic and literary uses are problematic, insofar as they must be said to be false.[7] Although they are false, they might not be forceless, and the notion of force could serve here to reinstate such uses within the theory and philosophy of language. Unfortunately, Austin does not make this move but follows the tradition in excluding such uses from his consideration.

To substitute to the notion of truth-value, Austin calls on the term

[7] Among other things, it is this incapacity for conception of language to account for literary statements that brought to the fore the notion of fiction, from Russell's bald king of France to Walton's focus on make-believe. The problem of fiction has gained crucial significance in the contemporary philosophical landscape, but poetry seems to place a limit to the notion of fiction as it blurs the fiction/nonfiction divide.

felicity. However, in describing felicitous and infelicitous performative utterances, he specifies that some utterances do not take part in this game of felicity. Those utterances are 'non-serious' uses of language:

> Secondly, as *utterances* our performatives are *also* heir to certain other kinds of ill which infect *all* utterances. And these likewise, though again they might be brought into a more general account, we are deliberately at present excluding. I mean, for example, the following: a performative utterance will, for example, be *in a peculiar way* hollow or void if said by an actor on the stage, or if introduced in a poem, or spoken in soliloquy. This applies in a similar manner to any and every utterance—a sea-change in special circumstances. Language in such circumstances is in special ways—intelligibly—used not seriously, but in ways *parasitic* upon its normal use—ways which fall under the doctrine of the *etiolations* of language. All this we are *excluding* from consideration. Our performative utterances, felicitous or not, are to be understood as issued in ordinary circumstances. (Austin 1975: 21–22)

From the outset, Austin considers poetic utterances 'hollow or void'. This view is not new to philosophy of language and reflects the traditional rejection of poetic uses as deviant uses. It is rather common in theorising language to distinguish between ordinary and poetic uses, the latter being deviances from the former. However, such a view fails to acknowledge the fact that 'deviant' uses might become ordinary. So-called 'dead' metaphors are an example of such bringing non-ordinary uses in ordinary language. More widely, everyday language is full of creative uses of language which outgrow so-called ordinary uses. By excluding poetic utterances from his conception of language, Austin limits its scope and makes a metaphysical move.[8]

Further in *How to Do Things with Words*, Austin makes a second distinction between 'the locutionary act (and within it the phonetic, the phatic, and the rhetic acts) which has *meaning*; the illocutionary act which has a certain *force* in saying something; the perlocutionary act

[8] The charge of making a metaphysical move is one Derrida raises against Austin in his article 'Signature, Event, Context' which gave rise to the famous debate with Searle. Searle's defence of Austin argues that the parasitic-ordinary distinction is not metaphysical but strategic. However, this defence misses Derrida's point which is precisely to say that if one posits a difference (be it only strategically), one commits to a system of metaphysical dualisms and its related hierarchies. Distinguishing from the outset ordinary from parasitic uses might be only strategic but it does not remain at this level. As soon as the distinction is posited, values are attributed to both terms of the dualism (traditionally, ordinary is good, parasitic bad). The whole idea of deconstruction is to undercut this system of metaphysical dualism by showing the fluidity of terms and values. Derrida's aim is therefore not to promote parasitic uses over ordinary ones, as this promotion would reaffirm the dualism, but to undercut it and consider the distinction impossible: for Derrida, one can never say whether a statement is ordinary or parasitic for those attributes can only be given within the game of metaphysical dualisms. Without entering in detail here, Derrida's view of Austin is in my opinion much more positive than what Searle's reply suggests: I believe Derrida considers Austin's theory as potentially germane but raises a reserve regarding the distinction he makes between ordinary and parasitic uses.

which is *the achieving of* certain *effects* by saying something' (Austin 1975: 121). As Austin mentions it, the distinction between illocutionary acts and perlocutionary acts 'seems the likeliest to give trouble' (Austin 1975: 110). In order to distinguish them, Austin shows that the illocutionary force is not a consequence of the locutionary act (in the sense of a physical consequence), but a convention. The illocutionary act might have an effect, but this effect is due to the conventions that regulate the use of language. The perlocutionary act is on the contrary not conventional as the use of language aims at certain effects without having those inscribed in a convention.

In this context, poetic utterances are considered 'hollow or void' from any illocutionary force, insofar as such utterances are not conventional. As Joe Friggieri argues, there is a 'suspension of illocutionary force' (Friggieri 2014: 58) in poetic utterances. However, two problems arise from such a conception: first, is this suspension always clearly stated and understandable? Most misunderstandings in ordinary speech might come from a misinterpretation of the status of the speech act. Second, is it possible to always distinguish an ordinary from an extraordinary use of language? Without entering into the details of Derrida's discussion of Austin, Austin's requirement for a 'total context' in order to distinguish ordinary from extraordinary uses is something Derrida considers impossible. Indeed, Austin considers that performative utterances, unlike statements, can be understood only by calling on 'the total situation in which the utterance is issued' (Austin 1975: 52). Is this situation fully determinable? Derrida argues against Austin by saying that the context is never completely determined, and that Austin's requirement is thus impossible to meet. If the context can never be determined as ordinary with complete certainty, it is impossible to distinguish ordinary from parasitic uses (Derrida 1988: 14).

In his various works, Stanley Cavell attempts to take poetry seriously in order, among other aims, to save Austin from Derrida's criticism. In *A Pitch of Philosophy*, he shows that the Derrida-Searle debate is the product of a mutual misunderstanding which has caused more trouble to philosophy than needed. Searle's criticism of Derrida indeed misses the point that Derrida sees something valuable in Austin, whereas Derrida puts too much focus on Austin's rejection of poetic language, which should not, according to Cavell, be interpreted as a rejection.[9] Cavell considers Austin's ordinary to be an opposition to both the metaphysical and the formal, not to the poetic or the literary. As he argues: 'That in literary studies Austin's ordinary language is instead thought to be contrasted with literary language means to me that Austin has not there been received' (Cavell 1994: 62). One way of receiving Austin in literary studies would therefore be to set aside the

[9] Recent scholarship has given much attention to the Derrida-Searle debate, as it marks an important moment in the history of the analytic-continental divide, see for instance Raoul Moati's and Jesus Navarro's book-length explorations of the debate (Navarro 2017; Moati 2014).

ordinary/literary distinction. Stanley Fish suggests such a reading of ordinary language: 'What philosophical semantics and the philosophy of speech acts are telling us is that ordinary language is extraordinary because at its heart is precisely that realm of values, intentions, and purposes which is often assumed to be the exclusive property of literature' (Fish 1982: 108). Following Cavell's and Fish's leads, recent trends in ordinary language philosophy have made of literature an important aspect of their investigation of language.[10] To that extent, Austin's project might be closer to Derrida than one initially thinks. According to Cavell, Derrida is wrong to consider Austin to be rejecting poetic language, as he argues that in fact Austin's theory can be of use for literary studies. In some sense, Cavell's whole philosophy is an attempt to take literature seriously from within the framework of ordinary language philosophy.[11]

We have seen that by depriving poetry from any performative and illocutionary force, Austin strips poetic language from any impact on the ordinary world. By placing such uses on a stage, far away from ordinary uses, he isolates poetry from ordinary language. In such a view, the value of poetry would only lie in an abstract and vain play on and with language: abstract because unrelated to any social and practical reality, vain because unable to affect the socio-political world. The forcelessness of poetry in language entails a forcelessness in the everyday world. Austin is not the only philosopher to put poetry in such a remote place, Sartre suggests something similar, following the acknowledgment that poetic language is remote from the ordinary.

Sartre argues that if a poetic utterance has no force, illocutionary and performative, it cannot have any political impact. He asks: 'How can one hope to provoke the indignation or the political enthusiasm of the reader when the very thing one does is to withdraw him from the human condition and invite him to consider with the eyes of God a language that has been turned inside out?' (Sartre 1988: 34). To place poetic language apart from ordinary language is to place poetry in no position to influence the everyday politicised world. This separation might seem surprising as, in *What Is Writing?*, Sartre defines literature in terms of political commitment, in terms of the effects literature can have in a political framework. This definition however concerns

[10] See Toril Moi's latest book, which questions the distinction between ordinary and literary language from within the framework of ordinary language philosophy: 'Ordinary language is certainly not the opposite of 'literary' language. (In my view, there is no such thing as 'literary language.') Nor is ordinary language the opposite of 'extraordinary language.' The extraordinary is at home in the ordinary. (We share perfectly ordinary criteria for when to apply the concept.) There is nothing extraordinary about the extraordinary' (Moi 2017: 162). See also Maximilian de Gaynesford's work on Austin and poetry (de Gaynesford 2011; 2009).

[11] On that topic, Cavell's concluding question in *The Claim of Reason* reveals the central role of literature for philosophy: 'Can philosophy become literature and still know itself?' (Cavell 1979: 496).

only literature (and here more specifically the novel) and not poetry. Indeed, Sartre uses this definition to distinguish literature from other artforms: 'No, we do not want to "commit" painting, sculpture, and music "too," or at least not in the same way' (Sartre 1988: 25). Poetry is an artform different from literature and therefore an artform whose main characteristic is not to be politically committed. The definition of poetry in terms of 'language that has been turned inside out' prevents poetry from having any political influence. For Sartre—as for Austin—poetry uses a language which steps outside the bounds of ordinary language, that is outside the bounds of a language that can influence the ordinary world.

In order to define literature in political terms, Sartre considers it to be a means for action:

> Thus, the prose-writer is a man who has chosen a certain method of secondary action which we may call action by disclosure. It is therefore permissible to ask him this second question: 'What aspect of the world do you want to disclose? What change do you want to bring into the world by this disclosure?' The 'committed' writer knows that words are action. He knows that to reveal is to change and that one can reveal only by planning to change. He has given up the impossible dream of giving an impartial picture of Society and the human condition. (Sartre 1988: 37)

If the language of poetry is remote from the everyday world, if it is as Sartre says a God's eye's view, literature is on the contrary rooted in the everyday world. The prose-writer uses language to influence the course of the world, to disclose something of the world and hence to change it. Words are action, words perform. For poets, however, language is something else: 'Poets are men who refuse to *utilize* language.' (Sartre 1988, 29) Sartre's definitions of poet and prose-writer however seem to rely on a contradiction. Indeed, is the poet not someone who 'has given up the impossible dream of giving an impartial picture of Society and the human condition' as well? For Sartre, the poet 'withdraw the reader from the human condition and invite him to consider with the eyes of God a language that has been turned inside out,' but does such a conception suggest that poetry has no effect, neither on the human condition, nor on language itself? If the poet is someone who changes language—who 'turns it inside out'—by using it (in opposition to maintaining language by utilising it), she might be someone who affects the human condition in greater ways than the prose-writer.

As long as we remain within the Austinian (and the philosopher of language) framework in which poetic statements are considered forceless, there is no way for the poet to affect the ordinary world. However, if we turn the Austinian distinction around and consider poetry forceful, the poet is then she who affects the human condition the most, insofar as she affects how language can be used. How can we turn this distinction around? The distinction Sartre suggests between prose-writers and poets as well as the distinction Austin establishes between ordinary and parasitic speech acts both rely on a conception of lan-

guage which postulate from the outset this distinction. It is by revaluating the conception of language without falling into the prejudice of the dualism between ordinary and extraordinary that we can give its force back to poetry.

2. *Nietzsche, Menke, and the notion of force*

If taking linguistic force away from poetry entails a political forcelessness, reinstating force in the language of poetry might make it relevant again for social and political concerns. The first step in giving poetry its force back is therefore to give it a linguistic force. Nietzsche's views on language offer useful insights in how force operates within language, and therefore how force can operate within poetic language (if there is even such a distinction between ordinary and poetic language). Claudia Crawford's reading of Nietzsche's theory of language provides an ideal starting point to explore the notion of force in Nietzsche's views on language:

> In the works of his last year another phase in Nietzsche's understanding of language is intensified and provides the material for a specific study. Language retains its effectiveness as force and play of forces, but now Nietzsche begins to lay more stress on the power which each individual instance of language production exerts as an instance of *value and action*. […] Language becomes a dynamic instance of interpretation and valuing, not in a critical sense of a subject who interprets values and then speaks or writes about those interpretations, but in a creative sense where the speaking or writing itself *is* the new value force embodied. (Crawford 1988: xiii)

This characterisation of Nietzsche's conception of language is Austinian in the sense that language is equated with action. If each instance of language is 'an instance of value and action', each instance could be read as a performative. Although Austin's initial limitation of performativity to a certain class of verbs seems to go against this idea, his later characterisation of utterances as all having a force goes in this direction. In this sense, Nietzsche shares Austin's idea that all speech acts are performative, i.e. have a force, but does not make the serious/non-serious distinction. For Nietzsche, all utterances are performative, including poetic and other non-serious ones.

This difference in their relation to poetry is the point of greatest dissent between Nietzsche's and Austin's conceptions of language. Indeed, while they share this common concern of language as power (a point of contact that Derrida already suggests[12]), the place attributed to poetry is radically different. Contrary to Austin's rejection of poetry, Nietzsche embraces poetic utterances by writing his philosophy in a

[12] Austin was obliged to free the analysis of the performative from the authority of the truth value, from the true/false opposition, at least in its classical form, and to substitute for it at times the value of force, of difference of force (illocutionary or perlocutionary force). (In this line of thought, which is nothing less than Nietzschean, this in particular strikes me as moving in the direction of Nietzsche himself, who often acknowledged a certain affinity for a vein of English thought.) (Derrida 1988: 13).

poetic way and even writing poems proper. His views on language, such as exposed in his early text *On Truth and Lie in a Nonmoral Sense* in which he considers language to be primarily metaphorical, bring to the fore the creative aspect of language and is at odds with Austin's more descriptive stance. This creative aspect is important because creating language, in the sense of naming, is a mark of power for Nietzsche, as he suggests in *On the Genealogy of Morality*[13]: 'the origin of language itself [is] a manifestation of the power of the rulers.' (*GM* I, 2) The relation between language and power lies in the fact that giving names, naming, is an act of power, of taking possession. Language does not only mirror the world in a neutral way but crafts it according to the will of the powerful. In *GM*, Nietzsche considers that the keys to shaping the world has been given to the rulers but, in *The Gay Science*[14], Nietzsche suggests that those who give names are those with originality: 'Those with originality have for the most part also assigned names.' (*GS* 261) This notion of originality brings us back to the realm of art and poetry. Reading this aphorism with *GM* in mind suggests that those who have power are not the rulers but the artists, those with force are those with originality.

As a shaping of the world, originality is a poetic force in the etymological sense of *poiesis*. It is a making of the world, which is, at the same time, an unmaking, as '[w]e can destroy only as creators.' (*GS* 58) This process of destruction and creation is primarily linguistic, as Nietzsche further argues: 'But let us not forget this either: it is enough to create new names and estimations and probabilities in order to create in the long run new "things."' (*GS* 58) The force of artists, and poets especially as they are primarily concerned with language, lies in their capacity to create new words and hence new things. The linguistic force of originality is both a destructive and creative force which modifies the world we live in.

Although the notion of force is hardly ever conceptualised as such in aesthetics, it is quite common to consider artworks to have a certain force, a certain effect. Nietzsche considers this force to be creative, and Christoph Menke offers a thorough exploration of the force of art in his book *Die Kraft der Kunst*. He constructs his notion of force in contrast to that of capacity, one being active and the other passive:

> Capacity makes us subjects who successfully take part in social practices, insofar as they reproduce their general form. In the play of *forces,* we are pre- and over-subjective agents who are no subjects; active, without self-consciousness; inventive, without aim.[15]

[13] (Nietzsche 2006) Hereafter *GM*.

[14] (Nietzsche 1974) Hereafter GS.

[15] My translation: '*Vermögen* machen uns zu Subjekten, die erfolgreich an sozialen Praktiken teilnehmen können, indem sie deren allgemeine Form reproduzieren. Im Spiel der *Kräfte* sind wir vor- und übersubjektiv—Agenten, die keine Subjekte sind; aktiv, ohne Selbstbewusstsein; erfinderisch, ohne Zweck' (Menke 2013: 13).

This opposition between passive capacity and active force is not without reminding Deleuze's reading of Nietzsche according to which there are active forces and reactive forces in Nietzsche's philosophy. Deleuze argues that positive forces are active, i.e. creative, whereas reactive forces are negative in the sense that they are always subject to a previous active force.[16] Similarly, Menke considers the notion of capacity to make us subjects, i.e. to submit us to a social practice, whereas the play of forces aims at freeing us from this subjection. In the play of forces, we are 'inventive without aim' because we are creative without being submitted to a conventional practice. In this sense, poetry might indeed not work within the realm of illocutionary forces, as those are conventional, but within perlocutionary forces. However, as we have seen with Austin, the distinction between illocutionary acts and perlocutionary acts 'seems the likeliest to give trouble' (Austin 1975: 110). This trouble might mean, and this is a path Nietzsche and Menke open, that this distinction should be abandoned.

More specifically, it might be impossible to distinguish the illocutionary from the perlocutionary. Perhaps, the perlocutionary can even become illocutionary in time. Although Austin does not discuss the ways in which the illocutionary force comes to existence (it seems to the contrary that for Austin this illocutionary force is either there or not, without any consideration about how it might appear or disappear), we could imagine that an unintended perlocutionary effect might, in time, repeat itself so regularly that it becomes a convention and thus an illocutionary effect rather than a perlocutionary one. A comparison with the theory of metaphor might enlighten this point. Indeed, it is common to distinguish dead metaphors from creative ones. A dead metaphor however is originally a creative one that has been so used that it does not appear as a metaphor anymore. Max Black compares conventional metaphorical uses such as 'cherry lips' to catachresis: 'So viewed, metaphor is a species of *catachresis*, which I shall define as the use of a word in some new sense in order to remedy a gap in the vocabulary. Catachresis is the putting of new senses into old words. But if a catachresis serves a genuine need, the new sense introduced will quickly become part of the *literal* sense' (Black 1955: 280). In this sense, a metaphorical sense can become literal. Once a metaphor becomes conventional, it loses its creative aspect. Similarly, once a perlocutionary

[16] 'The power of transformation, the Dionysian power, is the primary definition of activity. But each time we point out the nobility of action and its superiority to reaction in this way we must not forget that reaction also designates a type of force. It is simply that reactions cannot be grasped or scientifically understood as forces if they are not related to superior forces—forces of *another type*. The reactive is a primordial quality of force but one which can only be interpreted as such in relation to and on the basis of the active' (Deleuze 1983: 42). Even though Deleuze's interpretation of Nietzsche in terms of active and reactive forces is somewhat problematic, the distinction between active and reactive forces provides an insightful framework to consider how active forces are transformative, how the 'Dionysian power,' as Deleuze puts it, or the power of art, can actively transform the world by transforming values.

effect becomes conventional, it could become an illocutionary effect.

However, as conventions are always contextual and as the 'total context' of a speech act is, as Derrida argues, never fully determinable, these conventions are never completely determined. There is a lack of determinacy that appears more or less clearly in our uses of language (hence the possibility of misunderstanding or, in Austin's vocabulary, of misfire). It therefore seems that without this determination, the illocutionary force of a speech act is difficultly distinguishable from its perlocutionary force. In this sense, poetry contests the idea that the context and convention on which the illocutionary force relies can always be determined. In other words, there are so few determined contexts (if there are any) that the illocutionary force would only apply to very few utterances. Rather than suspending the illocutionary force, poetry reveals that it might not be the most important force in language. The illocutionary force would in this sense be a subcategory of the perlocutionary, without there being such a sharp distinction between them. So following Nietzsche and Menke, we must reassess Austin's distinction between illocutionary and perlocutionary, considering all utterances to have a force, relying to various degrees on conventions and contextual cues.

It remains however to be clarified what poetic speech acts do, what their effects are. Menke considers art to have a certain effect, to be a certain making which comes close to Nietzsche's views:

> What art makes is not an object of knowledge, because what art makes does not have its ground in knowledge. Hence philosophical aesthetics has called this making 'obscure' (Baumgarten): the aesthetic making is not a self-conscious activity, because there is no aesthetic making without the action of 'unconscious forces' (Herder). This action is play: the connection and disconnection and the new connection and again disconnection of images in the acts of imagination.[17]

There is a play of forces which connects and disconnects (in Nietzsche's terms: creates and destroys) images. In poetry, such images are words and the poets are those who connect and disconnect words, not only between one another, as in a sentence or spatially on the page, but also between language and the world, thus revealing how the world is made up with words. The effect of art, and more specifically of poetic speech acts, is to reveal the play of forces in which we are embedded. This play of forces is creative and hence also creates the agent. The poet is not subject to language, she is not in a reactive stance towards language, she exists even before this first determination. In poetics, the

[17] My translation: 'Das Machen der Kunst ist nicht Gegenstand des Wissens, weil das Machen der Kunst nicht im Wissen seinen Grund hat. Deshalb nennt die philosophische Ästhetik dieses Machen "dunkel" (Baumgarten): Das ästhetische Machen ist nicht eine selbstbewußte Handlung, denn es gibt kein ästhetisches Machen ohne das Wirken "unbewußter Kräfte" (Herder). Dieses Wirken ist Spiel: das Verbinden und Lösen und Neuverbinden und Wiederauflösen von Bildern in Akten der Einbildung' (Menke 2013: 67).

unconscious always plays a role, not in the sense that the originality or the genius of the poet lies within the psychoanalytic unconscious, but because 'the world of which we can become conscious is only,' as Nietzsche argues, 'a surface- and sign-world, a world that is made common and meaner.' (*GS* 354) The poet's play with the unconscious is therefore a broadening of the scope of language and hence an expansion of the world; it is an active force of transformation and creation.

We have seen that in the play of creative and destructive forces, the poet is a pre- and over-subjective agent who does not operate on the world of consciousness, but on that of the preconscious. In terms of language, the poet reveals the limits of linguistic conventions (illocutionary force) because they can never be fully determined. In other words, the poet arises as subject insofar as she brings the linguistic, social, and even poetic conventions into question. The notion of 'subject' is at the heart of Meschonnic's and Kristeva's conceptions of poetry and both reveal the importance of the transformative force(s) of poetry: transformation of language, of the subject, of society. In this sense, the force of poetry is a revolutionary one.

3. *The revolution of poetry*

We have seen that Nietzsche's views of language bring force back into poetic language, and that Menke's conception of art brings to the fore the idea that art is a creative and transformative force. It remains to be seen how this force operates in poetry, and I will focus on two theoretical works to do so, Meschonnic's *Celebration of Poetry* and Kristeva's *Revolution in Poetic Language*, and analyse one poem, Caroline Zekri's 'Un pur rapport grammatical'. Both Meschonnic and Kristeva consider poetry to be a transformative force: poetry does not leave the subject (as reader or writer) unchanged. She undergoes a transformative process due to a transformation of language. As Meschonnic clearly states: 'there is a poem only if a form of life transforms a form of language and if reciprocally a form of language transforms a form of life.'[18] This double transformation of a form of life and a form of language—both being intimately related to one another—is precisely where the force of poetry operates. Because of this transformation of her form of life, the reader or writer cannot maintain the same attitude within and towards her surrounding world. In taking poetry seriously (unlike Austin's rejection of poetic utterances in the realm of the 'non-serious'), she must accept this transformation of her form of life.

What is central to Meschonnic's views is the idea that poetry and life are intimately bound to one another. Against the views which argue that poetry is remote from the everyday politicised world—and hence remote from ordinary life—Meschonnic considers poetry and life to be

[18] My translation: 'il y a un poème seulement si une forme de vie transforme une forme de langage et si réciproquement une forme de langage transforme une forme de vie' (Meschonnic 2001: 292).

closely related as the title of his book *La rime et la vie* makes it explicit. The poem must therefore not be understood in terms of work, i.e. in an essentialist way, but in terms of activity. The poem as a work of art is a working on changing the world and the subject through the transformation of language. Against the idea that poems have a truth, Meschonnic argues that they have an activity, an effect, a force. With Meschonnic, we move from Austin's notion of illocutionary force to something closer to Menke's notion of force: poems have a force that build a subject rather than subject users to linguistic conventions. It is in this sense that poems can be revolutionary. Against the idea that poetry, as a form of literature, can teach us facts about the world, poetry teaches us a different way of seeing the world and its revolutionary force resides precisely in its capacity to bring us to see things differently.

If we take seriously the idea that poems transform a form of language and a form of life, thinking about poetry is not something which concerns only small details of our lives. Quite to the contrary as, according to Meschonnic, 'to think the poem, one must rethink the whole of language, and the whole relation between language, art, ethics, and politics.'[19] Insofar as poetry transforms our form of language and our form of life, it has an ethical and political impact. For Meschonnic, thinking poetry requires rethinking language and, through this reconceptualisation of language, rethinking our being in the world. If our ways of being in the world are dependent on our language, i.e. if our form of life is dependent on our form of language, and if a conception of language must account for poetry—because one can hardly argue that poetry is not related to language—we must modify our conceptions of language which fail to account for poetry and by changing those, change our ways of being in the world. In this sense, poetry becomes a topic of central philosophical significance. Meschonnic brings to the fore the political, ethical, and foremost existential dimensions of poetry, insofar as language itself bears these political, ethical, and existential dimensions.

The ways in which poems can affect our ways of being in the world are multiple, but recent evolutions in poetry show that, against the romantic ideal of the poet in her ivory tower, the contemporary poet, following Baudelaire whose 'halo slipped from [his] head, down onto the muddy street' (Baudelaire 2008: 91), must be in the 'muddy' everyday world and act within it. Some poets operate this move from the tower to the ground by focusing on the quotidian and the everyday as poetic material. Kenneth Goldsmith's transcription trilogy in which he transcribes weather reports (*Weather*, 2005), traffic (*Traffic*, 2007) or sports broadcast (*Sports*, 2008) participate in this idea of an 'uncreative writing' which casts a new light on everyday texts and situations. Fol-

[19] My translation: 'C'est pourquoi, pour penser la poésie, le poème, il y a à repenser tout le langage, et tout le rapport entre le langage, l'art, l'éthique et le politique' (Meschonnic 2001: 256).

lowing Moi's title, they operate a *Revolution of the Ordinary* in which the ordinary is both the subject and the object of the revolution: the ordinary is the material for the revolution and is what becomes changed through the poetic work (Moi 2017).

Franck Leibovici's work on poetic documents explores this transformative force of poetry through the way poetry modifies the use (and hence the meaning) of documents. Part of his analysis relies on Charles Reznikoff's *Testimony* which, according to Leibovici, transforms court transcripts in such a way. In the process, the court transcripts lose their documentary quality and become what Franck Leibovici calls 'poetic documents' that overcome the categories of true and false. As he argues in discussing Reznikoff's *Testimony*, the court transcripts are modified insofar as they become fictional, not in the sense that they become false, but that revealing their linguistic nature 'automatically suspends the categories of truth and false.'[20] What is central is not that these documents become poetic, but the realization that any document has the potential to become poetic and hence that the categories of 'document' and of 'poetry' are to be reassessed.[21] This reassessment of categories entail, as Meschonnic suggests, a rethinking of politics and ethics through the reconceptualization of language and art. Kenneth Goldsmith even considers that the simple reproduction of a text has sometimes more impact than a 'creative' poetic production: 'Sometimes, by the noninterventionist reproduction of texts, we can shed light on political issues in a more profound and illuminating way than we can by conventional critique' (Goldsmith 2011: 84). Such 'poetic documents' reveal the fact that poetry has the force to modify a form of language and hence a form of life, but also, following Meschonnic, that a form of life transforms a form of language. Poetry does not come out unchanged from its encounter with documents, quite to the contrary. It is crucial to note that according to Leibovici 'poetic documents' are not static, are not to be considered fixed entities as works of art, but are processes that can always evolve: 'the output, through the successive redescriptions, has gained a strong analytic power: at the same time process and product, small machine to redscribe and output of a redescription, it can be applied to some situations, working as a transportable poetic document, as if dematerialized' (Leibovici 2007: 68). Against the essentialization of the work of art, poetic documents are workings of art, they act upon the world by acting upon language. The apparent oxymoron 'poetic documents' reveals that the quotidian world affect poetry and that, in turn, poetry affects the quotidian world.

By affecting the quotidian world, poetic documents are not only con-

[20] 'par fiction, nous n'entendons pas quelque chose relevant de l'imaginaire, du "forgé de toutes pièces," mais simplement quelque chose exhibant sa nature langagière, suspendant ainsi automatiquement les catégories de vrai ou de faux (critères de vérité)' (Leibovici 2007: 35).

[21] Gaëlle Théval suggests that these poetic practices resemble readymades in art (Théval 2015).

ceptual devices (as the term conceptual poetry describing Goldsmith's work might suggest) but can also aspire to having a transformative and even political impact. An example of such a poetic document with political aspirations is Zekri's 'Un pur rapport grammatical' which combines excerpts from the 'Report of the Mapping Exercise'[22] from the Office of the High Commissioner for Human Rights (OHCHR), analysing violence in Democratic Republic of the Congo (DRC) between 1993 and 2003, and excerpts from individual evaluation forms from the '*Permanence d'accueil et d'orientation des mineurs isolés étrangers à Paris* (PAOMIE)' (Zekri 2015: 16), showing how French authorities evaluate under 18 migrants in Paris. The confrontation and reconfiguration of these documents reveal a certain use of language that the poetic document aims to disrupt. Zekri distinguishes both sources by using short quotations organised as verses from the Report and longer sentences from the evaluation forms. This distinction generates a contrast between a form of emotional violence in shorter verses and a form of rejection of emotion in longer sentences. There is something of a distinction between a more poetic structure on the one hand and a more narrative or argumentative structure on the other.

Here are the first lines of Zekri's poem:

on the pretext of searching their genitals for minerals
including diamonds, gold, copper, cobalt, cassiterite (tin ore) and coltan
are alleged to have mutilated and disembowelled a pregnant woman
stripped, manhandled and even severely beaten with nail-studded pieces of wood
for having worn trousers
and 17% of global production of rough diamonds
and two girls aged six and seven

When asked about his older sisters' age, he begins counting out loud. It is difficult to believe that he doesn't know his older sisters' age. He says he can't explain.[23]

[22] https://www.ohchr.org/Documents/Countries/CD/DRC_MAPPING_REPORT_FINAL_EN.pdf

[23] (Zekri 2015: 7) I used the translation from the report to translate the parts taken from there and translated myself the sentences taken from the evaluation cards:

sous prétexte de chercher des minerais dans leur parties génitales
dont les diamants, l'or, le cuivre, le cobalt, la cassitérite et le coltan
auraient mutilé et éventré une femme enceinte
dénudées, molestées et battues sévèrement avec des planches cloutées
pour avoir porté un pantalon
et 17% de la production mondiale de diamants bruts
ainsi que deux fillettes de 6 et 7 ans

Lorsqu'on lui demande l'âge de ses sœurs aînées, il commence à calculer à voix haute. Il est difficilement crédible qu'il ne sache pas l'âge de ses sœurs aînées. Il dit qu'il ne peut expliquer.

To understand the first lines of the poem, it is necessary to explain the title 'Un pur rapport grammatical' as it plays on (at least) three possible meanings of *rapport* in French.

1. The obvious meaning when we know the source of these sentences, *rapport* means report. That it is a grammatical report might mean that Zerki's poem focuses on the linguistic aspect of the report and plays with it. Indeed, except for the 'and' in verse 6, all these sentences can be found in the report from the OHCHR and the addition of 'and' is a grammatical one establishing a grammatical relation (this notion relation will be of importance in the third line of interpretation) between the two verses. We can understand that the reorganisation of the report operates on grammatical grounds. If we focus on the grammar of the sentence, two elements stand out.

First, in verse 3, the subject of the verb 'are' must grammatically be 'diamonds, gold, copper, etc.' from verse 2, which suggests that the violence is operated by the minerals themselves and, by metonymy, by the mining industry, something which is epitomised in the last two verses of the poem: 'some had their anuses ripped with a knife/ by multinationals' (Zekri 2015: 15). In this last verse, Zekri adds 'by' to generate once again a grammatical relation in a way similar to verse 6. The same play with grammar can be seen in verse 4 as the feminine plural of the adjectives ('stripped,' 'manhandled,' 'beaten') can only refer to genitals in verse 1, thus bringing the attention to the womanhood of the victims.

Second, in verses 5, 6, and 7, Zekri uses a zeugma to join three objects by using one verb, playing with three meanings of *porter*: 1. to wear trousers, 2. to carry diamonds (that relates to the first verse where minerals are hidden in genitals), 3. to carry a child (as in being pregnant). This zeugma therefore gives three reasons to explain the violence in verse 4 and places these reasons on a same level, considering them of equal importance.

2. *Rapport* means ratio. Although the idea of a grammatical ratio might be somewhat strange, the use of percentages in the report and the poem brings this aspect to the fore. The poem brings the reader's attention to the use of percentages and the depersonalisation that they operate. People become numbers, *ratio*, rather than victims. Furthermore, there is a relation between the importance of statistics in economy, in the mining industry in this case, and in the evaluation of the damages of the industry. The word 'ratio' therefore becomes a grammatical connector between the industry and its damages.

3. *Rapport* means relation. As already mentioned in point 1, the notion of relation is crucial. The poetic document establishes grammatical relations between aspects that are not necessar-

> ily so connected. Furthermore, the poetic document establishes a relation between the report and the evaluation forms, which has at least two effects. First, these evaluation forms show how people in Paris are evaluating migrants from their perspective without considering the effects of the violence that they have been through, hence the repetition of 'difficult to believe' in many of the excerpts.[24] Second, it establishes a relation between what has happened in the DRC and what happens in Paris: multinationals can go to DRC and commit violence, but migrants cannot come to Paris and have a supposedly incoherent story.

Zekri's poem therefore shows how poetry transforms an ordinary form of language (report and evaluation forms) and reveals something through this transformation. There is, if not a direct political claim, an injunction to discover and uncover relations (*rapport*) that operate in language. In this sense, I believe we can extend the conception of document even to some forms of poetry that do not use ready-made texts. Insofar as poetry uses language, and if there is an attention to the language used, i.e. if there is a linguistic or metapoetic dimension to the poem, all poems can to some extent be documentary. [25] If we consider that the language of poetry is not ontologically distinct from ordinary language, ordinary language becomes the document that the poetic phenomenon disrupts and transforms. It is in this sense that there is a *Revolution of the Ordinary* following Moi or, following Kristeva, a *Revolution in Poetic Language.*

Kristeva's title can be understood in two different ways, depending on the interpretation of the 'in'. First, it can be interpreted as a revolution occurring within the realm of poetic language and a large part of her work indeed investigates the changes that occur within the language of French poetry at the end of the 19th century. But this interpretation remains within the framework according to which there is such a thing as 'poetic language', even though it is subject to changes and revolution. Second, from a broader perspective, it can be interpreted as a revolution operated by poetic language. The revolution is not occurring within poetic language but is caused by poetic language. According to this second interpretation, poetry becomes a revolutionary force.

[24] 'difficile à croire' (p. 8), 'difficilement crédible' (p. 9), 'peu de crédibilité' (p. 11), 'peu crédible' (p. 12), 'peu crédible' (p. 14).

[25] This is especially true of a certain trend in contemporary French poetry, inspired by the American Objectivists among others, that takes the ordinary, the quotidian, the everyday as material for poetic explorations. In their anthology on Contemporary French poetry, *Writing the Real*, Nina Parish and Emma Wagstaff explore this 'refiguration of the everyday' among other trends (Parish and Wagstaff 2016). The quotidian has been of central importance in 20th century French poetry (Sheringham 2006), and this importance continues nowadays, especially in the journal *Nioques* and at the publisher 'Questions théoriques'.

We have seen with Meschonnic that poetry investigates the relation between forms of language and forms of life. It is precisely through this investigation that Kristeva considers poetry to be a revolutionary force: 'But mimesis and poetic language do more than engage in an intra-ideological debate; they question the very principle of the ideological because they unfold the *unicity* of the thetic (the precondition for meaning and signification) and prevent its theologization.' (Kristeva 1984: 61). Insofar as poetic truth is a way into the understanding of the relation between language and the world, and the constitution of this relation, it is not external to ideological, social, and political debates, but at their very heart. Poetic language is a space for transgression in which the foundations of the ordinary norm are fundamentally brought into question. Kristeva uses the term 'theology' to describe this relation to a given that is never put into question, which relates to Nietzsche's claim in *Twilight of the Idols*[26]: 'I'm afraid we are not rid of God because we still believe in grammar…' (*TI* 'Reason' 5).

Against the traditional view that posits poetic language as distinct from the ordinary, Kristeva considers it to lie at its very heart and to question the prejudices embedded in our uses of language. As an enemy within, poetic language becomes a revolutionary force. Indeed, whereas positing poetic language as outside only weakens its force, making it an external other which can easily be rejected, considering poetic language as lying at the heart of the ordinary gives it a force of changing things from within. For Kristeva, poetry affects language in such ways that our conceptions of the world cannot remain unchanged.

In order to conceptualise how poetry operates such a change, Kristeva elaborates the notion of practice and, as we have seen with Meschonnic, a poem is not a work but a working, an activity, a practice:

> The text thereby attains its essential dimension: it is a *practice* calling into question (symbolic and social) *finitudes* by proposing *new signifying devices.* In calling the text a practice we must not forget that it is a new practice, radically different from the mechanistic practice of a null and void, atomistic subject who refuses to acknowledge that he is a subject of language. Against such a 'practice,' the text as signifying practice points toward the possibility—which is a jouissance—of a *subject who speaks his being put in process/on trial through action.* In other words and conversely, the text restores to 'mute' practice the jouissance that constitutes it but which can only become jouissance through language. (Kristeva 1984: 210)

Kristeva's substitution of 'poem' by 'text' broadens the scope of what poetry is and can do. Shifting from poem to text undercuts all formal definitions of poetry and moves towards the notion of practice. Whereas the notion of poem remain within the framework of mechanistic practice—there are rules defining what a poem is and, in this sense, a poetics is a set of rules for literary creation[27]—Kristeva considers the prac-

[26] (Nietzsche 2008) Hereafter *TI*.

[27] Such a conception of poetics is at the heart of Aristotle's *Poetics* in which he famously gives the rules to evaluate a good tragedy. In this context, poetics is a way

tice of the text to be a revolutionary one 'calling into question (symbolic and social) *finitudes*.' There is a social aspect to textual practice insofar as the text operates a critical force attempting to overcome established social rules. However, like Nietzsche and Menke, Kristeva considers this critical aspect to be only one side of practice. Indeed, following Nietzsche's idea that we can destroy only as creators, the criticism of finitudes occurs only through the 'proposing [of] *new signifying devices*.' The text is therefore not only a place of criticism against established social values, but also a place of creation of new means of signification. The text therefore becomes a signifying practice in which the subject comes to jouissance, to living language in a positive way. The focus on the notions of text and practice leads Kristeva to consider poetry as combining two forces in language: a destructive and a creative one. It is only in combining these two forces that poetry can be revolutionary. If it focused only on the critical side, it would remain forceless because criticising language by using language is a self-contradiction. However, focusing on the creative side of language allows to propose new signifying practices which replace the old ways of thinking, hence destroying as creators.

Both Meschonnic and Kristeva bring our attention on the fact that poetry is revolutionary insofar as it is a practice aiming at modifying our ways of being in the world through the modification of language. Against the essentialisation of poems as works of art, their views consider necessary to approach poems as workings of art, as doing something to language and to the subject (as reader and writer), thus relating to Leibovici's idea that poetic documents are processes rather than fixed and stable entities. The force of poetry lies in this doing. Poetry therefore teaches us about the creative potential that lies within all forms of language. Whereas considering poetic language as essentially distinct from ordinary language makes poetic language forceless, bringing the poetic back within the ordinary reveals the creatives capacities of language, the transformative and revolutionary forces that animate our uses of

of classifying and evaluating artworks in respect to their relation to representation or *mimesis*: 'Now epic poetry and the making of tragedy, and also comedy and dithyrambic poetry, as well as most flue-playing and lyre-playing, are all as a whole just exactly imitations, but they are different from one another in three ways, for they differ either by making their imitations in different things, by imitating different things, or by imitating differently and not in the same way' (Aristotle 2006: 19). Jacques Rancière describes such a poetics as the poetic regime of art: 'I call this regime *poetic* in the sense that it identifies the arts—what the Classical Age would later call the 'fine arts'—within a classification of ways of doing and making, and it consequently defines proper ways of doing and making as well as means of assessing imitations' (Rancière 2004: 17). Against this poetic regime, Rancière consider the aesthetic regime which moves away from the Aristotelian classification in terms of representation towards an identification of art in its singularity. Kristeva's shift from poem to text is an attempt to move from the poetic regime (where the poem is classified as a poem) to the aesthetic regime (where the text exists in its singularity without any presupposed classification).

language. A philosophy of language that overlooks this creative force cannot account for the enormous potential that lies at the heart of language. Such a philosophy of language would be, as Kristeva argues at the very beginning of *Revolution in Poetic language*, 'nothing more than the thoughts of archivists, archaeologists, and necrophiliacs' (Kristeva 1984: 13). If the force of poetry is a revolutionary one, then, in respect to philosophy of language, poetry forces philosophy to operate such a revolution in order to account for the creative uses of language that are present in our everyday life. If, as Meschonnic argues, poetry does something to language, it also does something to philosophy of language: it reveals the shortcomings of conceptions of language and the necessity for any serious philosophy of language to account for poetic phenomena. The force of poetry therefore brings philosophy to rethink its categories (language, truth, fiction, ordinary) while modifying our ways of being in and of affecting the world. By doing things to words, poetry affects our conceptual scheme and our forms of life. If, following Austin, poetry cannot teach us how to do things *with* words because it cannot be taken seriously, it reveals us how to do things *to* words and, through this doing, how to transform and affect the world.

References

Aristotle. 2006. *Poetics*. Translated by Joe Sachs. Newburyport: Focus Publishing.

Austin, J. L. 1975. *How to Do Things with Words*. Edited by J. O. Urmson and Marina Sbisà. Cambridge: Harvard University Press.

Baudelaire, C. 2008. *Paris Spleen and La Fanfarlo*. Translated by Raymond N MacKenzie. Indianapolis: Hackett.

Black, M. 1955. 'Metaphor'. *Proceedings of the Aristotelian Society* 55 (1): 273–94.

Cavell, S. 1979. *The Claim of Reason Wittgenstein, Skepticism, Morality, and Tragedy*. Oxford: Clarendon Press.

———. 1994. *A Pitch of Philosophy: Autobiographical Exercises*. Cambridge: Harvard University Press.

Crawford, C. 1988. *The Beginning of Nietzsche's Theory of Language*. Berlin and New York: De Gruyter.

Deleuze, G. 1983. *Nietzche and Philosophy*. Translated by Hugh Tomlinson. New York: Columbia University Press.

Derrida, J. 1988. *Limited Inc*. Edited by Gerald Graff. Translated by Jeffrey Mehlman and Samuel Weber. Evanston: Northwestern University Press.

Fish, S. 1982. *Is There a Text in this Class?* Cambridge: Havard University Press.

Friggieri, J. 2014. 'Etiolations'. In B. Garvey (ed.). *J. L. Austin on Language*. Houndmills: Palgrave Macmillan: 50–69.

de Gaynesford, M. 2009. 'Incense and Insensibility: Austin on the "Non-Seriousness" of Poetry'. *Ratio* 22 (4): 464–85.

———. 2011. 'How Not To Do Things With Words: J. L. Austin on Poetry'. *The British Journal of Aesthetics* 51 (1): 31–49.

Gibson, J. (ed.) 2015. *The Philosophy of Poetry*. Oxford: Oxford University Press.

Goldsmith, K. 2011. *Uncreative Writing: Managing Language in the Digital Age*. New York: Columbia University Press.

Kristeva, J. 1984. *Revolution in Poetic Language*. Translated by Margaret Waller. New York: Columbia University Press.

Lakoff, G. and Johnson, M. 2003. *Metaphors We Live By*. Chicago and London: University of Chicago Press.

Lamarque, P. 2017. 'What Is the Philosophy of Poetry?' In S. Majetschak and A. Weiberg (eds.). *Contemporary Approaches to the Aesthetics of Nature and of Arts. Proceedings of the 39th International Wittgenstein Symposium in Kirchberg*. Berlin; Boston: De Gruyter: 109–26.

Leibovici, F. 2007. *Des documents poétiques*. Marseille: Al Dante//Questions théoriques.

Menke, C. 2013. *Die Kraft der Kunst*. Berlin: Suhrkamp.

Meschonnic, H. 2001. *Célébration de la poésie*. Paris: Verdier.

Moati, R. 2014. *Derrida, Searle: Deconstruction and Ordinary Language*. New York: Columbia University Press.

Moi, T. 2017. *Revolution of the Ordinary: Literary Studies after Wittgenstein, Austin, and Cavell*. Chicago and London: University of Chicago Press.

Navarro, J. 2017. *How to Do Philosophy with Words: Reflections on the Searle-Derrida Debate*.

Nietzsche, F. 1974. *The Gay Science: With a Prelude in Rhymes and an Appendix of Songs*. Translated by Walter Kaufmann. New York: Vintage.

———. 2006. *On the Genealogy of Morality and Other Writings: Revised Student Edition*. Edited by Keith Ansell-Pearson. Translated by Carol Diethe. Cambridge: Cambridge University Press.

———. 2008. *Twilight of the idols*. Translated by Duncan Large. Oxford: Oxford University Press.

Parish, Nina, and Emma Wagstaff, eds. 2016. *Writing the real: a bilingual anthology of contemporary French poetry*. London: Enitharmon Press.

Rancière, J. 2004. *The Politics of Aesthetics: The Distribution of the Sensible*. Translated by Gabriel Rockhill. London and New York: Continuum.

Rorty, Richard. 1967. *The Linguistic Turn Recent Essays in Philosophical Method*. Chicago: The University of Chicago Pr.

Sartre, J-P. 1988. *'What Is Literature?' And Other Essays*. Cambridge, Mass: Harvard University Press.

Sheringham, M. 2006. *Everyday Life: Theories and Practices from Surrealism to the Present*. Oxford: Oxford University Press.

Théval, G. 2015. *Poésies ready-made: XXe-XXIe SIÈCLE*. Paris: Editions L'Harmattan.

Wittgenstein, L. 1998. *Culture and Value*. Translated by Peter Winch. Oxford: Blackwell.

Zekri, C. 2015. 'Un Pur Rapport Grammatical'. *Nioques* 15: 5–16.

Croatian Journal of Philosophy
Vol. XXI, No. 61, 2021
https://doi.org/10.52685/cjp.21.1.8
Received: August 18, 2020
Accepted: December 18, 2020

A Literary Aesthetics of War Crime

RAFE MCGREGOR
Edge Hill University, Ormskirk, UK

In order to develop a literary aesthetics of war crime, I examine the phenomenon of moral immunity in military memoir. Using three paradigmatic examples of memoirs of unjust wars characterised by the routine perpetration of war crimes, I argue that moral immunity is achieved by means of three literary devices: literary irresponsibility, ethical peerage, and moral economy. I then employ the proposed literary aesthetics of war crime to provide an answer to the perennial question of the relationship between literature and morality as well as to two specific instantiations of this question, the value interaction debate in literary aesthetics and the ethics of reading in literary theory. My conclusion is that the literary aesthetics of war crime demonstrates both that there is a systematic relationship between aesthetic value and moral value and that there is no systematic relationship between literary ambiguity and moral uncertainty.

Keywords: Autobiography; colonialism; moral value; philosophy of literature; war.

1. *Introduction*

The value interaction debate and the ethical turn in criticism both address the perennial question of the relationship between art and morality, the first from the perspective of analytic aesthetics and the second from the perspective of literary theory. The value interaction debate has not been restricted to literature, although it has been primarily concerned with works of narrative art and, in consequence, literature and film (McGregor 2014: 450). As one might expect from the distinct approaches, literary aesthetics has focused on the specific issue of whether a moral defect in a work is (also) an aesthetic defect while literary theory has explored the ethics of reading, the relationship be-

tween literary responsiveness and ethical responsibility in the reception of texts, more broadly.[1] Where the relation between moral and literary defects has thus far been debated almost entirely in short form and almost exclusively restricted to fiction, the relationship between responsiveness and responsibility has been explored in both articles and monographs and included both fiction and nonfiction.[2] Tess McNulty (2018: 384) notes that although the value interaction debate and ethical turn in criticism were contemporaneous, the philosophical and critical movements have remained almost entirely independent of each other for two decades. She characterises the former as examining the relationship between aesthetic reception and ethical thinking in terms of the success (or failure) of works to produce their intended effects and the latter as examining that relationship in terms of the distinction between ambiguity and edification on the one hand and didacticism and indoctrination on the other hand.

McNulty proposes a rapprochement of literary aesthetics and literary theory by employing tools from the former to debunk a conclusion in the latter. She uses the analytic concepts of uptake, failed uptake, and imaginative resistance to argue that—contrary to received critical wisdom—there is no systematic relationship between either ambiguity and uncertainty or between didacticism and conformity. In other words, McNulty demonstrates that sophisticated ambiguity does not necessarily produce the moral uncertainty characteristic of the ethical knowledge conveyed by literature (and, similarly, that simplistic didacticism does not necessarily produce the conformity associated with literature as a vehicle for politics). There is no corresponding received aesthetic wisdom, although the recent history of the value interaction debate has been dominated by two theories, moderate moralism and robust immoralism. The former, with which Noël Carroll (1996: 236) initiated the debate, holds that a moral defect can be an aesthetic defect in a work of literature. The latter, introduced by A.W. Eaton (2012: 290), holds that a moral defect can be an aesthetic merit in a work of literature. While the positions appear compatible Eaton is clear that they are not, as they both argue for a systematic relation between specific kinds of moral defect and aesthetic value—reaching conclusions in opposite directions.

My approach is firmly anchored in literary aesthetics, but my method involves the extension of its analytic concepts to a practice or genre to which it is not usually applied, i.e. to nonfiction.[3] In this particular combination of approach and method, I take my lead from Sarah Worth

[1] For the purpose of this paper, I assume that there is a strong relation between aesthetic value and literary value (value *qua* literature), i.e. aesthetic value is at the very least a significant component of literary value.

[2] The two publications that initiated the respective debates are exemplary in this regard; see: Carroll (1996) and Miller (1987).

[3] For paradigmatic examples of these two approaches to fiction, see: Lamarque and Olsen (1994) and Friend (2012).

(2017: 39–67), who shifts the focus of the debate about the values of literature from its previous emphasis on the values of fiction to an examination of what is distinctive about nonfiction. One of the consequences of this shift is to prioritise the distinction between narrative and non-narrative representation over the distinction between fictional and nonfictional representation and while my literary aesthetics of war crime is not reliant on this priority, I do endorse it. In the course of her argument for the value of reading fiction, Worth (2017: 101) uses memoir as a case study of a blended genre: 'Memoir shares qualities of both fiction (because of its literary style, prose, character developments, and plot formations) and nonfiction (because it is all supposed to be true).' I shall also focus on memoir, specifically on military memoir. Where McNulty integrates literary aesthetics with literary theory by employing analytic concepts from the former to debunk a conclusion in the latter, I employ analytic concepts to life writing, a practice or genre that is standardly explored by means of literary theory. My proposal for a literary aesthetics of war crime will then provide a solution to both the value interaction debate and the ethics of reading, as well as a characterisation of the relationship between literature and morality more generally.

2. *Moral immunity*

The literal meaning of *biography* is life-writing and *life writing* is a term employed to describe practices that include, but are not restricted to, the writing of letters, diaries, memoirs, autobiographies, and biographies. Hermione Lee (2009: 5) uses the following definition of biography as her starting point: 'Biography is the story of a person told by someone else.' Linda Anderson (2011: 6–8) traces the first use of 'autobiography' to the end of the eighteenth century and notes that a hierarchy of self-representation was quickly established in which autobiographies were regarded as more valuable than memoirs in virtue of their greater seriousness, which was in turn a function of prioritising teleology over chronology. Laura Marcus (1994: 3) identifies the greater value associated with autobiographies over confessional writing and memoir on the basis that the former are '"sincere"', understood as exploring the totality of the self. She notes that autobiography is a fundamentally hybrid and unstable genre, bridging divides between subject and object, private and public, fiction and fact, and literature and history. This is reflected in the many varieties and experiments with autobiography that have occurred since the term was first used, from William Wordsworth's *The Prelude or, Growth of a Poet's Mind; An Autobiographical Poem* (1805), a blank verse poem of fourteen books, to J. M. Coetzee's *Summertime* (2009), an autobiography that combines fiction with fact and is written in the third person.

I delineate autobiographies as, paradigmatically: (1) narrative representations, (2) narrated in the first person, (3) with an identity of (a)

author, (b) narrator, and (c) protagonist, (4) all of whom are real people. Within this genre, I conceive of memoir as a subcategory that typically shares all four of the above characteristics and typically represents a part rather than the whole of the author's life, in consequence of which it may lack the seriousness (teleological) and sincerity (exploratory) associated with autobiography. John Gibson (2012: 109) describes autobiography as 'one of the last unexplored frontiers in literary aesthetics' and I concur with this assessment. I have selected military memoirs as an especially ethically problematic kind of memoir and selected three extreme cases, where the wars involved are unjust and where the authors are either active participants in or passive witnesses to war crimes. I do not offer a definition of unjust war, but take the colonial conflict in George Robert Elford's *Devil's Guard* (1971) and the white supremacist wars in Peter McAleese's *No Mean Soldier: The Autobiography of a Professional Fighting Man* (1993) to be uncontroversially unjust. I take the Iraq War in James Ashcroft's *Making a Killing: The Explosive Story of a Hired Gun in Iraq* (2006) unjust in virtue of the widespread scepticism concerning its legality (Chilcot 2016). I take war crimes to include any act in breach one of any of the four Geneva Conventions of 1949 or their three Additional Protocols of 1977 and 2005 (ICRC 2020).

Devil's Guard may be the most disclaimed book ever, beginning with a Publisher's Note that distances the publisher from the point of view expressed in the memoir and an Introduction in which the author, George Robert Elford (a pseudonym), abdicates from any moral responsibility for reproducing the story of the narrator and protagonist, Hans Josef Wagemueller (also a pseudonym), a junior officer in the *Waffen-SS* in the Second World War and in the *Légion étrangère* in the First Indochina War. The memoir is a curious and complex mélange of fiction and fact. On the one hand, the book has been widely denounced as Neo-Nazi fiction and its two sequels were published as war fiction (Elford 1988, 1991; Gibbons-Neff 2017). On the other hand, there is evidence both that the Foreign Legion allowed the many German volunteers to serve together and that fifty of the original members of the *36. Waffen-Grenadier-Division der SS* (Wagemueller's unit, thinly-disguised in the memoir) not only survived the war, but escaped prosecution for war crimes (Windrow 2018; Ingrao 2006). The memoir demonstrates an undeniable familiarity with the operational tactics of the 36th *Waffen-SS* Division and there are many passages whose authenticity is verified by more rigorous military histories (see: Fall 1961). As such, I shall treat the memoir in the same manner as the other two while recognising that some of the content is likely fictional.

No Mean Soldier is closer to an autobiography than a memoir, beginning with Peter McAleese's birth in 1942 and chronicling the next fifty years of his life. I shall focus on his military service for two white supremacist regimes from 1977 to 1983, for Ian Smith's Rhode-

sian Front government during the Rhodesian Bush War and then for P. W. Botha's National Party government during the South African Border War (chapters six to eight of twelve). McAleese is credited as the author 'with Mark Bles' and copyright is shared by the two. As the use of his name disqualifies Bles from being a ghostwriter, I take the relationship between Bles and McAleese to be similar to that between Elford and Wagemueller (and Ashcroft and Thurlow below), i.e. the former was responsible for the form of the work and the latter for its content. *No Mean Soldier* should not be confused with *Beyond No Mean Soldier: The Explosive Recollection of A Special Forces Operator!* (McAleese 2015), which is actually a second edition of the autobiography. McAleese joined the Parachute Regiment of the British Army at the age of seventeen and became one of the youngest men to pass selection for the Special Air Service (SAS) a mere two years later. He was dismissed from the SAS twice for drunken brawling, completed his military service in 1969, and received three custodial sentences for intimate partner violence over the next seven years. He began a career as mercenary in 1976, fighting with the infamous Colonel Callan in the Angolan War of Independence, as a foreign volunteer in the Rhodesian Security Forces and the South African Army, and as a mercenary again in Colombia from 1988 to 1989, where he was contracted to assassinate Pablo Escobar.

Making a Killing is a memoir of James Ashcroft's (a pseudonym) work as a contractor for a UK-based private military company (PMC) called Spartan (a disguised name) from September 2003 to March 2005. He is the product of Winchester School and the University of Oxford and served in the British Army as an infantry officer from 1992 to 1998. The narrative begins with Ashcroft in his mid-thirties, giving up a career in law in the City of London to satisfy his longing to return to the adventure of military life. In Baghdad, Ashcroft joins a team of five other mercenaries, all ex-soldiers in their forties: Seamus and Les, former senior non-commissioned officers in the British Army; and Hendriks, Kobus, and Etienne, former special forces operators in the South African Army. He completes three contracts with the team over the next eighteen months: providing personal security for foreign journalists, escorting oil tankers from Kuwait to Baghdad, and training a security force to protect the city's water supply. Ashcroft and his professional collaborator, Clifford Thurlow, published a sequel, *Escape from Baghdad: First Time Was For the Money, This Time It's Personal*, in 2009. Unlike the prequel, *Escape from Baghdad* is lacking in authenticity and I take the pair of books to follow the pattern of the *Devil's Guard* series, i.e. whatever the relationship between narrative and history in the first, it is substantially looser in the second.

In spite of—or perhaps, more accurately, in consequence of—the criminal wars and war crimes represented in each of these memoirs, the authors (by which I refer to Wagemueller, McAleese, and Ashcroft

respectively) seek some kind of acquittal, amnesty, or absolution from their readers. This can be understood as a straightforward desire to avoid being judged as morally abhorrent, which seems a minimal requirement if the reader is to accept the author's invitation to the read the narrative *qua* memoir, i.e. with an interest in the life of the author as opposed to a historical interest in the wars represented or a criminological interest in the crimes represented. I shall call this authorial intention, that reader engagement with the work is characterised by some degree of identification, empathy, or sympathy, *moral immunity*. In the remainder of this paper, I argue that moral immunity is achieved by the deployment of three distinct devices: literary irresponsibility, ethical peerage, and moral economy. In literary aesthetic terminology, these devices are all aesthetic merits employed in a morally defective manner, i.e. inventive design features intended to diffuse moral responsibility. In literary critical terminology, the devices are all sophisticated instantiations of ambiguity employed to a simplistic didactic end, i.e. inventive uses of language intended to justify white (or Western) supremacism. McNulty (2018: 384) notes that sophisticated ambiguity has standardly been associated with modernist or canonical texts so its perhaps surprising appearance in these memoirs evinces Gibson's claim about the lack of philosophical exploration of autobiography.

3. *Literary irresponsibility*

In 'Fiction and the Nonfiction Novel', Peter Lamarque (2014: 83–84) establishes a taxonomy of the relationship between fiction and nonfiction in the novel. His taxonomy identifies six categories of novel, including paradigmatic fiction, paradigmatic nonfiction, and the nonfiction novel. Paradigmatic nonfiction includes biographies, autobiographies, and memoirs, which may be more or less literary in the relationship between their form and content. 'Nonfiction novel' was coined by Truman Capote to describe his *In Cold Blood: A True Account of a Multiple Murder and Its Consequences* (1966) and presents nonfictional content in the form of a novel, i.e. using the literary techniques associated with paradigmatic fiction in order to increase the work's impact on readers (Lamarque 2014: 88). Lamarque's (2014: 98) objection to nonfiction novels is the moral irresponsibility of the authors:

> The "nonfiction novelist" cannot have it both ways: cannot claim the high ground of the serious reporter or historian and also the imaginativeness and inventiveness of the novelist. There is a moral dimension here as well as a theoretical tension between practices.

Lamarque criticises authors such as Capote, Norman Mailer, and Tom Wolfe for writing realist novels in the tradition of Tobias Smollett, Charles Dickens, and Evelyn Waugh (paradigmatic fiction), but claiming the accuracy of the representation of reality associated with biography, autobiography, and memoir (paradigmatic nonfiction). This

critique can be extended beyond the specific criticism of writing one type of novel in the guise of another to a more general literary device, towards which Lamarque (2014: 103) gestures in his summary of his argument:

> where authors deliberately conceal their intentions, apparently inviting response under one set of norms (e.g., serious reporting) while hiding behind the privileges of another (e.g. literary licence), then, as we saw, the question of moral responsibility inevitably comes up.

I shall refer to the deliberate misrepresentation of literary content for the purpose of inviting response under an inappropriate set of norms as *literary irresponsibility*.

Devil's Guard is a eulogy for National Socialism, an apologia for colonialism, and a polemic for the escalation of United States' involvement in the Second Indochina War. By the time of its publication, the waging of that war was two years into its Vietnamization phase, an exit strategy that was justified on the basis of the Army of the Republic of Vietnam being strong enough to win the war on its own (Wiest 2002: 50–51). This context is significant as there is a much closer relationship between the military and political aspects of *Devil's Guard* than the other two examples, which for the most part focus on the trials, tensions, and thrills of contact with the enemy. The polemical intention behind the novel is literally rather than figuratively staged in the twelfth chapter (of eighteen), titled 'Dialogue with an Agitator' (Elford 1971: 221). Wagemueller describes a raid on a *Viêt Minh* village in which sixteen insurgents have been found drunk and incapable. Following the battalion's standard operating procedure, the Legionnaires execute the insurgents, bayoneting them to death while they are asleep to avoid wasting ammunition. When a Chinese commissar and his *Viêt Minh* deputy are captured while trying to escape, Wagemueller decides to take the advice of one of his officers, Erich Schulze, who suggests that a public debate with the prisoners will provide an entertaining interlude from the routine of operations. The bulk of the chapter (eleven of fourteen pages) comprises a discussion in which Schulze and his comrade, Bernard Eisner, engage the two commissars in a superficial (but nonetheless serious) elenchus aimed at proving the political, ethical, and economic superiority of capitalism over communism. The Chinese commissar, Kwang, is allowed to expound on the merits of communism at length and is even provided with a drink of water to facilitate further pontification. Schulze brings the debate to a close once he feels he has refuted Kwang's argument and honours his word by releasing both commissars from custody.

While the debate is in progress, Wagemueller reflects on its impact (Elford 1971: 232):

> The villagers listened in utter silence; their faces betrayed no emotion—only alertness. Some of the elder men were listening so intently that their mouths hung open and their eyes appeared transfixed on Schulze. I was

> not sure if all that Erich [Schulze] said had reached the people, and if so, how deeply his words had penetrated into their simple minds. I was sure of only one thing, that never before had they witnessed someone challenging the Viet Minh platform openly, in front of people. They had never heard someone denouncing the holiest of the Communist prophets and everything they stood for. No one could ever call a Viet Minh leader a liar and live to tell the story. Besides, no French officer in Indochina had ever bothered to talk to *les sauvages* on equal terms, and certainly not about political or economic issues.

There are several unintended ironies in this passage, not the least of which is Wagemueller's criticism of his French comrades for referring to the Vietnamese as *savages* mere sentences after he himself has referred to them as *simple*. While the debate provides a curious interlude in a narrative that is otherwise almost completely focused on combat operations, it is important to recognise its significance within Wagemueller's narration. When the Legionnaires learn that the *Viêt Minh* subsequently executed thirty people in the village, Wagemueller speculates that it was because they had repeated Schulze's argument. For all the incongruity of the debate, it constitutes an explicit justification of the German presence in Vietnam. The incident is, in fact, a clear case of literary irresponsibility because notwithstanding the bookending of the debate between two sets of war crimes (the execution of the insurgents and the execution of the villagers), the dialogue is represented as a genuine and legitimate exploration of two opposing views. Indeed, as the debate progresses, it is easy to forget that the commissars are being interrogated at gunpoint, uttering what they believe to be their last words.

A recurring rhetorical strategy in military memoir is the use of humour to represent situations that are either terrifying, tragic, or both. This of course reflects the way in which soldiers and non-combatants deal with repeated exposure to the threat of death, permanent injury, and capture. The representation of terror or tragedy in a humorous manner deploys literary irresponsibility when the humour is used to characterise the events rather than the author's response to those events, i.e. when the reader is invited to find the events themselves rather than the author's response amusing. McAleese makes occasional use of humour in *No Mean Soldier*, particularly in the two chapters that deal with his service in the Rhodesian Security Forces. The most extended example of literary irresponsibility is his narration of an incident at the school in Bindura, a mining town in north-eastern Zimbabwe (1984) with a population of approximately eighteen thousand, at the end of 1979. Over a thousand white Rhodesians, all in formal evening dress and heavily-armed, descend on the school hall for the annual ball. Once the festivities are underway, two Zimbabwe African National Liberation Army (ZANLA) insurgents make a token attack, firing a few rounds from their assault rifles. In response, drunken

guests charge from the hall, pouring fire into the darkness, causing the insurgents to withdraw. At the sound of a firefight, however, the Guard Force (heavily-armed African security guards), launch a counter-attack on what they assume are the insurgents, also firing into the darkness. The battle between the revellers and their guards lasts for hours, with hundreds of rounds fired. McAleese's (1993: 149) final observation is: 'Bindura looked as though it had been attacked by stormtroopers, with bullet strike on walls, smashed windows, shattered door, and riddled cars, but, once again, with the luck of the devil, no one was hurt.' The invitation to find the incident amusing belies the fact that it took place in a densely-populated urban area in a country that had already suffered fifteen years of civil war.

In *Making a Killing*, Ashcroft discusses both his personal motivation for joining Spartan and the political motivation behind the Coalition Force (CF) invasion of Iraq. I discuss the former in the next section. With respect to the latter, he is sceptical about the justification provided by the Bush and Blair administrations as well as the way in which the subsequent occupation is conducted by the Coalition Provision Authority (CPA), critical of both phases of the Iraq War (2003–2011), the invasion (March to May 2003) and the insurgency (2003–2011). PMCs such as Spartan were involved in the second phase and it was this phase that had the most devastating effect on the civilian population of Iraq: reliable estimates put civilian casualties during the invasion at approximately three thousand seven hundred and fifty, but those during the insurgency at approximately one hundred and twelve thousand four hundred and fifty (Conetta 2003; IBC 2020). Ashcroft is well aware of the dire management of the occupation, but nonetheless unequivocally frames the role of the PMCs in Iraq as serving the Iraqi population rather than the CPA. The following two statements are explicit, the first made in the first chapter (of twenty-six) and the second on the penultimate page of the Afterword:

> (1) 'We were *not* an occupying force safeguarding the second largest oil reserves on the planet. We were rebuilding Iraq to bring security to the Iraqis' (Ashcroft 2006: 6).
>
> (2) 'The water purification plants I protected prevented outbreaks of typhoid, cholera and gastroenteritis and have directly saved the lives of thousands of Muslim children, the sick and the elderly. PSDs [personal security details] don't get paid to kill people. We get paid to save lives and stay out of trouble' (Ashcroft 2006: 268).

In all three examples (Wagemueller, McAleese, and Ashcroft), literary content (interrogation at gunpoint, danger of civilian casualties, and sustaining the occupation) is deliberately misrepresented (as political debate, a hilarious near-miss, and public service) so as to invite response under a more morally-acceptable set of norms. The deployment of literary irresponsibility contributes to authorial moral immunity.

4. *Ethical peerage*

In the course of her explanation of rational response to ethical disagreement in 'Literature and Disagreement', Eileen John (2014: 239) contrasts epistemic peers with ethical peers. Using literary examples as evidence, she argues that conciliation is a rational response to an ethical peer, even when that peer is an epistemic inferior. Epistemic peerage is determined by a combination of possessing the relevant information and the relevant capabilities and the idea is that if two epistemic peers disagree, each has rational grounds for conciliation. In the epistemic case, conciliation is contrasted with steadfastness and involves an individual either withholding or lowering confidence in his or her initial belief. Ethical peerage is determined by a combination of possessing a capacity for ethical judgement and a comparable interest in the resolution of the disagreement. In the ethical case, conciliation is 'a matter of people revising their reasoning, or seeing the need to revise their reasoning, so that even if beliefs remain in place, whether and how those beliefs are grasped and supported has shifted in a conciliatory way' (John 2014: 240). The concept of an ethical peer is thus both more simple and more complex than that of an epistemic peer. It is simpler because all persons are assumed to be ethical subjects, i.e. to have the capacity for ethical judgement and moral responsibility. It is more complex because it presupposes the existence of an ethical community, which consists of ethical peers whose interests should be acknowledged and for whom the ethical judgement should make sense. Significantly, the status of ethical peerage is, unlike that of epistemic peerage, determined by a person's situation rather than their competence. In the relationship between reader and author, there is an initial assumption of *ethical peerage* and this assumption promotes conciliation rather than steadfastness in cases where the reader disagrees with the author's ethical judgement.

McAleese begins *No Mean Soldier* with two attempts to establish himself and his readers as peers in an ethical community, in order to gain the trust of his audience prior to disclosing his participation in criminal wars and perpetration of war crimes. The front matter of *No Mean Soldier* consists of a half title, title page, copyright, epigraph, contents, acknowledgements, and preface. The epigraph (which is very likely the first part of the book that will be read) and the preface (which is very brief, less than a single page) play a significant role in immediately and effectively inaugurating a relationship of ethical peerage between author and reader. I quote the epigraph, which is a quote from Theodore Roosevelt (quoted in McAleese 1993: v), in full:

> It is not the critic who counts, not the man who points out how the strong man stumbled, or where the doer of the deed could have done better. The credit belongs to the man who is actually in the arena, whose face is marred by dust and sweat and blood, who strives valiantly, who errs and comes short again and again, because there is no effort without error and short-

> coming, who does actually strive to do the deeds, who knows the great enthusiasms and spends himself in a worthy cause, who at best, knows in the end the triumph of high achievement, and at worst, if he fails, at least fails while daring greatly. His place shall never be with those cold and timid souls who knew neither victory nor defeat.

The first three paragraphs of the preface are (McAleese 1993: 1):

> Before you start, I want to clear up a couple of points. I've always been a professional operational soldier and not, repeat not, a "career" soldier of the sort who wants only to keep his nose clean and worries about his promotion and pension prospects. In fact, I don't think I made a bad depot soldier when necessary, and maybe my turnout has been smarter than most. I wonder how many soldiers nowadays bother to iron creases into their combat uniforms? I've never soldiered just for profit. During all my service in three regular armies, my pay was unimpressive by modern standards. I went from one fighting zone to another but I receive no pension. Nor do I have funds from another source. I've done it for the adventure, because I have always been a professional soldier, and because I love a fight. I've never been happier than in action.

In the context of the memoir, the epigraph identifies the man (and in memoirs such as this it is always a *man*) of action as having greater moral significance than the critic. This greater moral significance is afforded in virtue of the value of action as opposed to inaction, of imperfect strife in service of a worthy cause. In other words, it is morally better to perpetrate some immorality in the course of striving for a moral end than to perpetrate no immorality because one does not pursue that end in the first place. McAleese is of course implying that he is the man of action and the reader the critic. Though the wars of white supremacism in which he fought hardly constitute a worthy cause they are represented as wars against communism, as part of the Cold War in which the majority of McAleese's readership at the time of publication would likely have been sympathetic to the West. One might in fact take this epigraph to be not merely establishing McAleese as an ethical peer in spite of the criminal wars and war crimes, but to be establishing him as an ethical superior, above and beyond the judgement of anyone who is not a man of action, who is a *cold and timid soul*. The epigraph serves to introduce the preface, in which McAleese explicitly establishes himself as the man of action (a category that excludes career soldiers). He addresses the reader both directly and in a direct manner, inviting reception as a frank, sincere, and most importantly honest interlocutor. As such, the two parts of the front matter of the memoir constitute a statement of and an appeal for ethical peerage.

Like McAleese, Ashcroft takes pains to portray himself as a scrupulously honest narrator. He (2006: 3) is candid about the financial rewards of his work as a mercenary, stating at the very beginning of the narrative: 'We were in Iraq for the $500 a day we earned.' The financial rewards are always described as only one of two reasons, however, the other being his longing for a return to the adventure of military life.

Ashcroft reflects on both the complexity of his motivation as well as the changes to that motivation at several places in the memoir. At the beginning of chapter two, he (2006: 31) makes a direct appeal for ethical peerage in a manner similar to McAleese, addressing what he correctly imagines will be the main moral reservation for readers, the fact that mercenaries are literally paid to kill:

> Civvies often ask if you enjoy killing people. They assume killing someone means wandering along the high street and slaughtering an innocent passby with a loving family at home. But it's not like that. The people I end up killing are always in the act of actively trying to kill me in some murderous, violent and agonising fashion. So, no, I don't enjoy killing people, but, yes, I feel great afterwards because I feel the initial and immediate exhilaration at realising that I am alive and that the man who tried to kill me has failed.

This deployment of ethical peerage at the individual level is of course matched by the deployment of literary irresponsibility at the organisational level (discussed in the previous section) such that Ashcroft represents himself as seeking dangerous, well-paid work for a humanitarian cause rather than killing for money or killing for pleasure.

Wagemuller is by far the most arrogant and unremorseful of the three authors and narrates with the bombast, hyperbole, and braggadocio of a very vainglorious yarn-spinner. Notwithstanding his attitude and the extreme violence in which the narrative revels, he is careful to establish some kind of ethical community between himself and his readership. The first paragraph of the Foreword (the first part of the book narrated by Wagemueller rather than Elford) contains the following passage (Elford 1971: 12):

> I was a *kopfjaeger*—"headhunter," as our comrades of the Wehrmacht used to call us. We were a special task force of the *Waffen* SS—the "fighting SS"—which had nothing to do with concentration camps, deportations, or the extermination of European Jewry. Personally I never believed that the Jews could or ever would become a menace to Germany and I hated no people, not even the enemy. I never believed in German domination of the world but I did believe that Germany needed lebensraum. It was also my conviction that Communism should be destroyed while still in its cradle. If my beliefs should be called "Nazism," then I was indeed a Nazi and I still am.

There are two points of particular interest here. First, Wagemueller immediately distances himself from what, in the West at any rate, is usually considered the most morally abhorrent aspect of National Socialism, the Holocaust perpetrated against Europe's Jewish population. Second, he frames his Nazism in terms of anti-Communism, in much the same way as McAleese frames his own support of white supremacist regimes. For Wagemueller this has an additional significance in that, as noted in the previous section, the book is part polemic against a US withdrawal from Vietnam. By distancing himself from what many readers will regard as the worst National Socialist atrocities and aligning himself with anti-Communist sentiment, he makes a subtle (albeit typically arrogant) appeal to ethical peerage. In all three

examples, the authors (McAleese, Ashcroft, and Wagemueller) make explicit claims to ethical peerage with their readers, promoting conciliation in the face of apparent immorality (fighting for pleasure, killing for money, and National Socialism) and in so doing invite immunity from that immorality.

5. *Moral economy*

Norbert Götz (2015: 149) traces the first use of 'moral economy' to 1729 and notes the various denotations and connotations between then and historian E.P. Thomson's distinction between political economy and moral economy in 1971. The denotation with which I am concerned is both subsequent and specific, Carroll's (2013: 235) use of the term in explaining why many viewers have a pro-attitude to the fictional character Tony Soprano in the television series *The Sopranos* (1999–2007) when they would abhor his counterpart in reality. Carroll uses *moral economy* to describe the array of characters represented in a narrative that correspond and contrast with respect to their morality. Within the moral economy of *The Sopranos*, where most characters demonstrate significant ethical defects, Tony is one of the least reprehensible. The moral economy frames viewer engagement with the series and produces the pro-attitude in the following way (Carroll 2013: 243):

> the pro-attitude that we extend to Tony Soprano is a result of the fact that we are allied to him. And we are allied to him because in the fictional world of *The Sopranos* alternative alliances would either be worse morally or irrelevant.

Carroll claims that the distinction between viewer responses to Tony and to his real-world counterpart is a consequence of the extent of the ethical defects represented in the fictional world rather than any essentially fictional feature of the series. Moral economies are employed in both fiction and nonfiction.

Although Ashcroft recognises the complexity of his own motivations for joining Spartan and displays an ambivalent attitude towards both the CF invasion and the CPA occupation, he establishes a definitive moral economy of the insurgency. His (2006: 6) initial assessment of Iraq in 2003 is as 'on the slippery slope to chaos', which is—with the benefit of hindsight—unfortunately completely accurate. Ashcroft uses 'terrorists' (2006: 6) to describe his enemy, deliberately conflating 'Al Qaeda and radical Islamists' (2006: 7) with '"fedayeen"' (2006: 19) insurgents, all of whom are represented as exacerbating the chaos. The CF for whom the mercenaries are working, which in Ashcroft's case is the US Army, is represented as for the most part well-intentioned, but poorly-led and badly-trained, in consequence of which they often involuntarily or unknowingly contribute to the chaos. Finally, the private security contractors—particularly those working for Spartan—are framed as experienced professionals, the only faction whose intervention always reduces the chaos. The distinction between the Americans

and the contractors is particularly effective in that it not only distances Ashcroft from the extensive civilian casualties of the insurgency, but also—somewhat ironically considering they are only there for the money—sets the contractors up as morally superior to the soldiers. While Ashcroft is often complimentary about individual US soldiers, he makes numerous criticisms of their lack of professionalism throughout the memoir, most of which are aimed at their lack of concern for civilian casualties. The following three quotes are indicative:

(1) 'the trigger-happy Americans at the CF checkpoint' (Ashcroft 2006: 17).

(2) 'I remembered reading that for every 15,000 rounds of ammo the US military fires there is one fatality. This guy was doing his best to lower the average' (Ashcroft 2006: 28).

(3) 'The highest scoring killer of private security contractors up until then was, of course, the United States Army, seconded by terrorists, but only when catching stray terrorist fire because they were driving along in traffic mingled with a US patrol' (Ashcroft 2006: 71).

Ashcroft's moral economy clearly establishes the PMCs as the least morally reprehensible faction amidst the chaos and he makes a subtle but explicit claim to the moral superiority of the British military over the US military by comparing the rules of engagement in Northern Ireland with those in Iraq. Ashcroft has experience of the former as a British Army officer and experience of the latter as a private military contractor, which affords his view a certain authority. The most damning comparison is not, however between Northern Ireland and Iraq, but within Iraq, specifically the different strategies pursued by the British and US military components of the CF (Ashcroft 2006: 70–71):

> Unlike the Brits mounting occupation and peacekeeping duties, the US troops in Iraq, especially in Baghdad in late 2003 and through 2004, were the same guys who fought their way in. The poor sods in the 3rd Infantry Division had a combat mindset not in any way conducive to peacekeeping. As for their anti-ambush drills, they had to be seen to be believed. Every weapon in the convoy unloaded in a 360° arc into anything that moved… dogs, donkeys, children, buses, private contractors, you name it, got some.

In my discussion of the literary irresponsibility deployed by Ashcroft, I described how he framed the role of the PMCs as serving the interests of the civilian population rather than the CPA and the deployment of a moral economy of the insurgency both reinforces and is reinforced by the literary irresponsibility. The PMCs had nothing to with the invasion, whose justification is suspect, and do not dictate the terms of the occupation, which is being poorly-managed; instead, they use their superior skills and mindsets to ease the burden of the population by protecting journalists, petrol convoys, and the water supply. The moral economy and literary irresponsibility work in tandem to conceal the fact that the PMCs were essential to the maintenance of the occupation, the prolonging of the insurgency, and—in consequence—the suffering of the civilian population.

Wagemueller establishes a simple moral economy between the *Viêt*

Minh and the Legion in which the atrocities of the former are always prior to and more extensive than the atrocities of the latter. Although the moral economy of atrocity is a straightforward *us versus them*, it is at times deployed in a subtle manner, for example in the two incidents that bookend the political debate I discussed previously: the Legionnaires bayonet sixteen sleeping insurgents to death, but the *Viêt Minh* execute thirty villagers. The latter atrocity is, in numerical terms alone, almost twice as bad as the former. In the chapter that follows, Wagemueller (Elford 1971: 238–239) states: 'only the French "crimes" received blaring headlines. The Viet Minh atrocities (far more numerous and excessive) were given a few back-page lines once in a while.' The *Viêt Minh* atrocities are unequivocally *far more numerous and excessive* than the French. The second aspect of the moral economy of atrocity is perhaps more effective as it seeks to absolve the Legionnaires from—or at the very least reduce their moral responsibility for—the atrocities. The representation of French atrocities as a response to *Viêt Minh* atrocities (with an implied causal relation) is reiterated at several places in the narrative, and the following two are representative:

(1) 'I was there at the beginning and I know that it was not the French who started the atrocities and what one may rightly call genocide. Genocide is a Communist specialty' (Elford 1971: 70).

(2) 'We were not any better than the Viet Minh and we knew it. But we did want to fight a clean war and we were not the ones who started the atrocities. We only retaliated in kind. We could do nothing else' (Elford 1971: 127).

The perspective on the atrocities is that they are both a response to the Communist atrocities and a necessary evil.

McAleese justifies his active support of white supremacist governments by means of establishing a moral economy that is similar to, albeit distinct from, Wagemueller's. Where Wagemueller's moral economy of atrocity is used to represent the atrocities the Legion commits as being morally acceptable in virtue of being less extensive and a necessary response to the atrocities of the enemy, McAleese's moral economy of settler colonialism is deployed as evidence for what constitutes perhaps the most common apologia for colonisation, that it is in the interests of the colonised. McAleese uses the crimes of the military and civilian organisations fighting for black liberation to justify the continued existence of the two white supremacist governments, Ian Smith's Rhodesian Front and P.W. Botha's National Party. I take an example from each:

(1) 'By the time I arrived, both Robert Mugabe's ZANLA and Nkomo's ZIPRA had carried out countless attacks on the local population and the blacks suffered more than the whites' (McAleese 1993: 92).

(2) 'Ondjiva was a wreck by now, with few civilians left who could stand the constant battering and looting by FAPLA and SWAPO' (McAleese 1993: 171).

The first quote refers to the situation in Rhodesia in 1977 and the ZANLA and Zimbabwe People's Revolutionary Army insurgents. The second quote refers to the situation in Angola in late 1981. Ondjiva—a town near the South West African (now Namibian) border—was occupied by the South African Defence Force to use as a forward operating base against the People's Armed Forces of Liberation of Angola and the military wing of the South West African People's Organisation. In both cases, it is made clear that white supremacist rule is safer and more secure for African civilians, inviting the reader to ally herself with the colonial settlers rather than the black insurgents, as the lesser of the two evils. In all three examples, the authors (Ashcroft, Wagemueller, and McAleese) establish moral economies (of insurgency, atrocity, and settler colonialism) that frame the author's faction as the least morally reprehensible of the various options and thus invite moral immunity from their membership of those factions.

6. *Conclusion*

In the previous three sections I introduced three literary devices, Lamarque's literary irresponsibility, John's ethical peerage, and Carroll's moral economy. I showed that all three of these devices were deployed in all three of the military memoirs for the purpose of moral immunity, i.e. for creating some kind of identification, empathy, or sympathy with the authors in spite of their participation in criminal wars and war crimes. In *Devil's Guard*, Wagemueller represents an interrogation as a political debate, appeals to the anti-communist sentiments of his readers, and justifies his atrocities as a necessary response to his enemy. In *No Mean Soldier*, McAleese conceals tragedy by means of comedy, represents himself as a man of action pursuing a worthy end, and frames settler colonialism as being in the interests of the African population. In *Making a Killing*, Ashcroft represents the PMCs as serving the Iraqi population, himself as seeking adventure in a humanitarian cause, and the PMCs as the only faction reducing the chaos in occupied Iraq. The three devices employed together in each memoir create a complex narrative framework in which readers are invited to treat the author with what Carroll calls a *pro-attitude*—a vague, but useful term that includes a broad range of responses from actually identifying with the author to simply not regarding him as morally abhorrent. This invitation to adopt a pro-attitude to the author will not of course be accepted by all readers, but the invitation itself evinces coherence and precision in its design and construction. The use of literary devices in service of moral immunity is thus very similar to—if not identical with—what one would expect to find in literary works or canonical texts. I am not suggesting that these memoirs have equivalent literary value to, for example, the realist novels Lamarque mentions, but that the literary

devices they deploy should be taken seriously. I am, in other words, suggesting that the relationships among literary irresponsibility, ethical peerage, moral economy, and moral immunity constitute a literary aesthetics of war crime.

In the introduction, I described the literary aesthetic concern with the relationship between literature and morality as being focused on the value interaction debate, the question of whether a moral defect in a work is (also) an aesthetic defect. To date, the most popular answers are from Carroll, who argues for a systematic relation between moral defects and aesthetic defects, and Eaton, who argues for a systematic relation between moral defects and aesthetic merits. The literary devices I have explored in this paper are all, in analytic aesthetic terms, aesthetic merits, understood as adding value to the memoirs when they are being judged as works of literature. Moral immunity—i.e., some kind of acquittal, amnesty, or absolution from participating in criminal wars, war crimes, or both—is clearly a moral defect. The authors' deployment of the literary devices for the purposes of moral immunity is thus evidence of a systematic relation between aesthetic merits and moral defects rather than evidence of a systematic relation between aesthetic merits and moral merits. I also described the literary critical concern with the relationship between literature and morality in the introduction, characterised as being focused on the relationship between literary responsiveness and ethical responsibility in the reception of texts. McNulty identifies literary theory as associating sophisticated ambiguity with moral uncertainty and simplistic didacticism with political conformity. The literary devices I have explored are all, in literary critical terminology, sophisticated instantiations of ambiguity employed for the purpose of political conformity, i.e. creating moral immunity from fighting for colonial or neocolonial powers. In consequence, they offer evidence for McNulty's debunking of critical wisdom on the univocal relationship between literary ambiguity and moral uncertainty.

With this in mind, I consider the literary aesthetics of war crime proposed in this paper to provide a compelling (if not conclusive) solution to both the value interaction debate (favouring Eaton's robust immoralism) and the ethics of reading (favouring McNulty's debunking). Setting aside these specifically disciplinary concerns, there remains the perennial question of the relationship between literature and morality. In focusing on the way in which the authors of military memoirs use distinctively literary devices to achieve distinctively (im)moral responses in their readers, the literary aesthetics of war crime demonstrates that the moral dimension of literature cannot and should not be ignored. In arguing for the way in which literary irresponsibility, ethical peerage, and moral economy are deployed in the service of moral immunity I have focused exclusively on immorality in literature. I make no claims about the way in which different literary devices are deployed in the service of moral ends such as compassion, respect, and

justice. The evidence from immorality is nonetheless sufficient to show that literature should not be experienced, interpreted, or appreciated in isolation from its moral dimension. Whether one's preference in articulating literature is as an institution or a canon, the value of literary works or canonical texts for art, for culture, and for humanity is inextricably bound up with their moral value.[4]

References

Anderson, L. 2011. *Autobiography*. 2nd ed. Abingdon: Routledge.

Ashcroft, J. with Thurlow, C. 2006/2007. *Making a Killing: The Explosive Story of a Hired Gun in Iraq*. London: Virgin.

———. 2009. *Escape from Baghdad: First Time Was For the Money, This Time It's Personal*. London: Virgin.

Carroll, N. 1996. "Moderate Moralism." *British Journal of Aesthetics* 36 (3): 226–238.

———. 2013. "Sympathy for Soprano. " In N. Carroll. *Minerva's Night Out: Philosophy, Popular Culture, and Moving Pictures*. Chichester: Blackwell Publishing: 234–246.

Chilcot, J. (2016). *The Report of the Iraq Inquiry*. 6 July. Available at: <https://webarchive.nationalarchives.gov.uk/20171123122743/http://www.iraqinquiry.org.uk/the-report/>.

Conetta, C. 2003. "The Wages of War: Iraqi Combatant and Noncombatant Fatalities in the 2003 Conflict. Project on Defense Alternatives." *Research Monograph* 8 (20 October). Available at: <http://www.comw.org/pda/0310rm8.html>.

Eaton, A. W. 2012. "Robust Immoralism." *The Journal of Aesthetics and Art Criticism* 70 (3): 281–292.

Elford, G. R. 1971/1985. *Devil's Guard*. New York: Dell.

———. 1988. *Devil's Guard II: Recall to Inferno*. New York: Dell.

———.1991. *Devil's Guard III: Unconditional Warfare*. London: Hodder & Stoughton.

Fall, B. 1961. *Street Without Joy: The French Debacle in Indochina*. Mechanicsburg: Stackpole.

Friend, S. 2012. "Fiction as a Genre." *Proceedings of the Aristotelian Society* 112 (2): 179–209.

Gibbons-Neff, T. 2017. "This is the Nazi-apologia war novel that apparently inspired SEAL Team 6 combat tactics." *The Washington Post*. 10 January. Available at: <https://www.washingtonpost.com/news/checkpoint/wp/2017/01/10/this-is-the-nazi-apologia-war-novel-that-apparently-inspired-seal-team-6-combat-tactics/?utm_term=.64433f23e5db>.

Gibson, J. 2012. "Selves on Selves: The Philosophical Significance of Autobiography." *Journal of Aesthetic Education* 46 (4): 109–119.

Götz, N. (2015). "'Moral economy': its conceptual history and analytical prospects." *Journal of Global Ethics* 11 (2): 147–162.

Ingrao, C. 2006/2011. *The SS Dirlewanger Brigade: The History of the Black Hunters*. Trans. P. Green. New York: Skyhorse.

[4] I would like to thank Eileen John, Robert Stecker, and Vladimir Rizov for their invaluable assistance with this paper.

International Committee of the Red Cross 2020. *Geneva Conventions of 1949 and Additional Protocols, and their Commentaries*. Available at: <https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/vwTreaties1949.xsp>.

Iraq Body Count 2020. *Database*. Available at: <https://www.iraqbodycount.org/database/>.

John. E. 2014. "Literature and Disagreement." *Proceedings of the Aristotelian Society* 114 (3): 239–260.

Lamarque, P. 2014. "Fiction and the Nonfiction Novel." In P. Lamarque. *The Opacity of Narrative*. London: Rowman & Littlefield International: 83–104.

Lamarque, P. and Olsen, S. H. 1994. *Truth, Fiction, and Literature: A Philosophical Perspective*. Oxford: Oxford University Press.

Lee, H. 2009. *Biography: A Very Short Introduction*. Oxford: Oxford University Press.

McAleese, P. with Bles, M. 1993. *No Mean Soldier: The Autobiography of a Professional Fighting Man*. London: Orion.

———. 2015. *Beyond No Mean Soldier: The Explosive Recollections of a Former Special Forces Operator!* Birmingham: Helion.

McGregor, R. 2014. "A Critique of the Value Interaction Debate." *British Journal of Aesthetics* 54 (4): 449–466.

McNulty, T. 2018. "Literary Ethics, Revisited: An Analytic Approach to the Reading Process." *New Literary History* 49 (3): 383–401.

Marcus, L. 1994. *Auto/biographical discourses: Criticism, theory, practice*. Manchester: Manchester University Press.

Miller, J. H. 1987. *The Ethics of Reading: Kant, de Man, Eliot, Trollope, James, and Benjamin*. New York: Columbia University Press.

Windrow, M. 2018. *French Foreign Légionnaire versus Viet Minh Insurgent: North Vietnam 1948–52*. Oxford: Osprey.

Worth, S. E. 2017. *In Defense of Reading*. London: Rowman & Littlefield International.

Zimbabwe 1984. *1982 Population Census: A Preliminary Assessment*. Harare: Central Statistical Office.

Croatian Journal of Philosophy
Vol. XXI, No. 61, 2021
https://doi.org/10.52685/cjp.21.1.9
Received: August 2, 2020
Accepted: December 21, 2020

*Art and the Impossible**

BORAN BERČIĆ**
University of Rijeka, Rijeka, Croatia

In this article author contrasts possibilism (the view that art is about the logically possible and that it cannot be about the impossible) with impossibilism (the view that art can be and sometimes is about the logically impossible as well). Author argues in favor of possibilism. The main insight is that since impossible objects are necessarily non-existent art cannot be about them, it has to be about something that can exist. Also, author formulates five more detailed views about the issue. Further, author discusses related notions like imaginability and conceivability. Author holds that Hume's insight that an object cannot be conceived as non-existent counts in favour of possibilism. Besides, author introduces the distinction between real and apparent content of the work of art, believing that this distinction can be relevant in the discussion between possibilism and impossibilism. In the rest of the article author analyzes several prima facie counterexamples to possibilism (Jean-Luc Picard, Anna Karenina, paradox of patricide, Escher's graphics) and tries to explain them away.

Keywords: Art; object; content; representation; possibility; contradiction.

1. *Aboutness*

Works of art are typically and usually *about something*. In the case of a novel or a film it seems perfectly right to ask *What is it about?* Her-

* It is questionable whether determinate article can be used when we talk about necessarily non-existent things, after all, things that cannot be individuated.

** This work has been fully supported by the University of Rijeka, research project Uniri-human-18–239. I wish to thank to colleagues and friends whose comments and suggestions made substantial contribution to this work (in chronological order): David Davies, Nikica Petković, Takashi Yagisawa, Aleksandar Mijatović, Graham Priest, Iris Vidmar, Branko Kukurin, Dajan Plačković, Karlo Mikić.

man Melville's *Moby Dick* is about a whale hunter who is obsessed with a huge white whale. Martin Scorsese's *Taxi Driver* is about a young lonely veteran who decides to fight the filth of the big city. Film can be about the corrupt policeman or about the middle age crisis. They have their *subject matter*. They are *intentional*. Paintings and sculptures are also usually *about something*. Michelangelo's *David* depicts beauty and symmetry of a human body. Pablo Picasso's *Guernica* shows horrors of an air raid. Even pieces of music can represent something, although typically they don't. Bedrich Smetana's *Vltava* shows how small mountain brook grows into a big river. Miles Davis' *So What* captures the sound of traffic in a big city.

2. *Art and the Possible*

Sometimes works of art are about the *actual* things, people that really lived and events that really happened. Central characters of Tolstoy's *War and Peace* were not fictional. Napoleon and Kutuzov really existed.[1] Dostoevsky's *Crime and Punishment* was inspired by an actually existing man. The town of Guernica was really bombed. *Serpico* existed and really fought corruption in the NYPD. However, works of art are usually about *fictional* characters and events. Characters from novels and films typically *never existed*. But it does not matter whether they existed or not. They *could have existed*. There could have been police officers like *Kojak* or *Columbo*, just as there could have been women like *Emma Bovary* or *Anna Karenina*. Even if they are not *actual*, they are *possible*. This was the view of Aristotle. Talking about the difference between the historian and the poet, he says:

> It is not the function of the poet to relate what has happened, but what may happen—what is possible according to the law of probability or necessity. The poet and the historian differ not by writing in verse or in prose. … The true difference is that one relates what has happened, the other what may happen. (Aristotle, *Poetics*, Section I, Part IX)

The claim that art is about the possible might be misleading.[2] Although fictional characters are *possible*, art tells us something about the *actual* world. Ultimately, it tells us something about *ourselves*. After all, this is why it is important. If we focus on the philosophical debate about the ontological status of the fictional characters (which is a very interesting question)[3] we might forget why art really matters and why it is impor-

[1] Though, Russian formalists, for instance, insisted on the difference between literary character and real person. No matter of the degree of similarity, the two should not be conflated. Kutuzov from Tolstoy's *War and Peace* is a fictional character while real Kutuzov is not.

[2] For instance, it is not clear whether accepting this view commits us to the existence of possible objects. In my opinion, it does not. To say that x is possible is not to say that there exists possible object x. But this a matter of further debate.

[3] Very good overview of different accounts of the ontological status of fictional characters can be found in the first chapter of Thomasson's *Fiction and Metapysics* from 1999.

tant to *us*. One might even claim that fictional characters are nothing but the means for saying something about the actual ones. John Slesinger's *Midnight Cowboy* is about something that could happen to *any* young man coming from the countryside to the big city. *Anna Karenina* is about something that could happen to *any* married woman. So, they are about the *possible* courses of events that might happen to the *actual* people. Shakespeare's character *Shylock* from *The Merchant of Venice* obviously is a fictional character. But it is more than that. It is an incarnation of greed. And greed is an *actual* character trait of really existing people. Alienation felt by character *Meursault* in Camus' *The Stranger* is felt by many *actual* people. Old man *Santiago* from Hemingway's *Old Man and the Sea* is a fictional character but his will to prove himself is something that we all have. *Serpico* did really exist while *Dirty Harry* did not. Nevertheless, *Dirty Harry* is about the way in which many *actual* people would love to react. So, although the art is about the possible people and events, there is an important sense in which the art is about the *actual*.[4] Perhaps this is most important sense. When we talk about the work of art, a good answer to the question *What is it about?* has to have several layers.

3. *Art and the Impossible*

I hope that these comments help clarify the sense in which art is about the possible, that is, the sense in which fictional characters and plots could have existed. However, it seems that art is not only about the possible. It seems that it can be about the *impossible* as well. Some works of art seem to be about the impossible courses of events. Some novels, films, graphics, etc. *prima facie* represent objects and events that *cannot exist*. But how can that be? How could anything be about the impossible? How can anything be about the things that cannot exist?[5] We watch such movies, we like them, we understand plots, we understand stories, we love heroes and hate bad guys, … However, if these plots are not possible, what are these movies really about?

[4] Pierre Corneille's *Le Cid*, performed in Paris in 1636, was criticized on grounds that it was *unlikely* that Chimene marry Rodrigue (Le Cid) who killed her father. In classicism probability (*vraisemblance*) was seen as requirement of poetics. Possibility was not sufficient, characters had to be probable. This was the main issue in the *Querelle du Cid*—famous debate about the norms of literature.

[5] Graham Priest wrote a short story about *an empty box that contains something* (Priest 1997). In the story he and his friend move furniture. The box first was on the shelf, than they moved it to some other place, etc. In fact, there is nothing impossible in the story. The box is just an ordinary object, only its *name* is contradictory. In Dubravko Matakovič's cartoons one of the main characters is *barren mother*. Since she is the mother of the main character, she is obviously not barren. This oxymoron just adds to the overall absurdity of characters and plots, it is not constitutive for the plot. For this reason it would be wrong to say that these stories are about impossible objects. They are not about the impossible objects. Their authors do not describe something impossible. They only use oxymorons to increase the impression of absurdity or to prove a philosophical point.

On the one hand, it seems obvious that art cannot be about the people and events that cannot exist, because in that case it would be about nothing. But, as we saw, art is not about nothing. Art is about something. It has some content. It has some object that it talks about. Therefore, art cannot be about the impossible. The impossible does not and cannot exist. If something purports to be about the impossible, it cannot be about anything at all.[6]

However, art quite often depicts states of affairs that are at least *prima facie* impossible. People cannot fly but we read cartoons and watch movies about *Superman*, *Batman*, *Spiderman*, etc. Animals cannot talk but we enjoy *Aesop's fables*. Moreover, we believe that they have pedagogical and ethical value. We know that time-travel is not possible but it is not an obstacle for watching *Terminator*. We know that humans cannot become something else but we perfectly well understand episode from *Odyssey* when Circe turns Odysseus' sailors into pigs, or the episode from the *Bible* when Lot's wife is turned into a pillar of salt, or Kafka's *Metamorphosis* where Gregor Samsa becomes a bug. We know that nothing can move faster than light but we have no problems understanding the idea of *warp-drive* in *Star Trek*, even the idea of *transwarp*. We know that there are no magic wands and no magic words, nevertheless *Harry Potter* is very popular. We have fun watching *Body-swap* movies or *Adams Family*. We read *Gothic novels*. Borges, Marquez, and other writers from Latin America developed *Magic realism*. We know that we cannot walk in a circle always ascending or always descending and just because of that we admire Escher's *Ascending & Descending*. Many things from Dali's and Chagall's paintings cannot exist but we do not regard that as a failure.

One might object that I am doing two different things in this article: metaphysics of fiction and philosophy of literary criticism. Discussion about the ontological status of fictional characters belongs to the metaphysics of fiction, while distinction between the apparent and real content belongs to the philosophy of literary criticism. These two things might be related but one should not blur the distinction and lump them together. However, it is completely natural that the two come together. They cannot be separated. The first discussion directly and necessarily leads to the second one. We can reconstruct this path in four steps:

1) *A* has to be about something.
2) Apparently *A* is about X.

[6] Can we make an exhibition of the impossible objects? No, because impossible objects do not exist and can not exist. There would simply be nothing to exhibit. The gallery hall would be empty. Of course, we can make exhibition with the title *Perpetuum mobile* and expose a number of drawings from the patent office. But we could not expose *Perpetuum mobile* itself—a machine that produces energy without consuming it. The same holds for concepts like *faster than light*, *time travel*, *round square*, etc. The museum of impossible objects would be empty by necessity. Of course, museum of non-actual but possible objects would also be empty, but not necessarily. If actualized, these objects could be exhibited. On the other hand, impossible objects could not be exhibited under no circumstances because they could not be actualized.

3) But *A* cannot be about *X* because *X* is not possible.
4) Therefore *A* has to be about something else.

If the apparent content is logically impossibile then piece of art has to be about something else and we have to interpret it. We have to find out what is it really about. For this reason metaphysics of fiction and interpretation of fiction naturally come together, they cannot and should not be separated. How can we correctly estimate the ontic commitment of a novel if we do not know what is novel really about? Although interpretation in literature and in other branches of art can be very fishy, general principle seems obvious: no correct metaphysical analysis without correct interpretation! Also, no correct interpretation without correct metaphysical assumptions. How can we say what is novel really about if we are not aware of the relevant metaphysical and logical limitations?

Jan Alber asks "how readers can make sense of" unnatural narratives (Alber 2016: 3). I hope that this article may be of some help in answering that question.

4. *The Nature of Possibility*

It seems that we can explain away many of the above mentioned examples. There is a sense in which many of these cases are possible. Of course, the question is how restrictive or how permissive our criterion of possibility is. Criterion that is usual in philosophical discussions is very permissive, it is the criterion of *logical possibility*: whatever is not a contradiction is possible. As Hume said:

> 'Tis an establish'd maxim in metaphysics, *That whatever the mind clearly conceives, includes the idea of possible existence*, or in other words, *that nothing we imagine is absolutely impossible*. We can form the idea of a golden mountain, and from thence conclude that such a mountain may actually exist. We can form no idea of a mountain without a valley, and therefore regard it as impossible. (Hume, *Treatise*, Book I, Part II, Section II)

The idea is simple, elements that can be combined without contradiction give a possible state of affairs. Flying + man = flying man, fox + talking = talking fox, time + travel = time travel, horn + one = unicorn, etc. So, since there is no contradiction in the idea of a flying man, *Superman* and *Batman* are possible. Since there is no contradiction in the idea of a magic wand, *Circe* and *Harry Potter* are possible. Although we have no slightest idea about the underlying causal mechanism, it is possible that words like *Abracadabra* or *Expelliarmus* have causal effects. Since radical skeptical scenarios are not contradictory, *Matrix* is possible. Since Cartesian insight that we are essentially thinking things and only accidentally have bodies is not a contradiction, *Freaky Friday* and other *body swap* movies are possible, etc.

Logical possibility is very permissive criterion.[7] Here is well known passage from Plantinga:

[7] Some accounts of possibility are much more restrictive. See, for instance, *A Combinatorial Theory of Possibility* of D. M. Armstrong and Brian Skyrms

> I think Socrates could have been an alligator; for I think he could have had an alligator body. At least he could have had an alligator body during part of his career. We have no difficulty in understanding Kafka's story about the man who wakes up one morning to discover that he has the body of a beetle; and in fact the state of affairs depicted there is entirely possible. In the same way it is possible that I should awaken one morning and discover (to my considerable chagrin) that my body had been exchanged for an alligator body. (Plantinga 1978: 65)

However, even if we accept so permissive criterion of possibility, this still does not mean that all the examples can be successfully explained away. It seems that some cases stubbornly resist: encounters of past and future selves, encounters of actual and possible selves,[8] time travel paradoxes,[9] temporal loops,[10] teleportation,[11] graphics of the impossible objects, nonexistent protagonists, incoherent stories, living deads, barren mothers, empty boxes that contain something, etc. These are cases that are at least *prima facie* logically impossible. This article is about such cases. The question is whether works of art can be about such cases. *Can art be about the logically impossible?* Or, talking in terms of possible worlds, the question is whether art can be about the logically impossible worlds.

In contemporary fertile and interesting discussion about the *unnatural narratives*, all impossibilities are treated together, as a single phenomenon. Jan Alber in *Unnatural Narrative* from 2016 deals with "physically, logically, or humanly impossible scenarios and events that challenge our real-world knowledge" (Alber 2016: 3) However, in philosophy there is a deeply entrenched distinction between *factual possibility* and *logical possibility*. Here factual possibility encompasses biological, technical, physical, and other kinds of possibilities that we know through experience. Logical possibility is conceptual possibility and we know it *a priori*. Personally, I am not very fond of this distinction, but there is a good reason for accepting it. Physically impossible objects are *conceivable*, while logically impossible objects are *incon-*

(Armstrong 1989) or *Modal Realism* of Mondadori and Morton (Mondadori and Morton 1976) not to be confused with Modal Realism of David Lewis.

[8] David Lewis argues that travel between possible worlds is not possible. (Lewis, 1986, 80) Saul Kripke says that there is no telescope for seeing different possible worlds (Kripke 1980: 44). Nevertheless, we enjoy movies where people encounter their possible selves.

[9] Ted Sider points to the fact that *Terminator 2* is incoherent (Sider 2009: 309). Nevertheless, we watch it. Though, Sider believes that time travel is not necessarily paradoxical.

[10] D. H. Mellor argues that temporal loops are necessarily impossible because in that case events should cause themselves and that is inconsistent (Mellor 1998: 132). His argument is strengthened with te fact that he relies on the causal theory of time. But if this is so, then the question is what is film *Groundhog Day* really about.

[11] Derek Parfit argues that teleportation is not a process of travelling but rather process that annihilates the original person and creates a replica somewhere else (Parfit 1984: 200).

ceivable. This difference is of paramount importance, especially in the philosophical analysis of art. Even if the events and objects depicted in the works of art are factually impossible, they are conceivable! We can conceive of flying men and talking animals, and in such cases we understand what is the work of art about. On the other hand, we cannot conceive logically impossible situations. If the piece of art is declared to be about round square or wooden iron, we cannot know what it is about because we cannot conceive of such objects.

5. *Imaginability and Existence*

There is another characteristic of imaginability and conceivability which is relevant in this context. We cannot imagine an object as nonexistent. If we imagine it at all, we imagine it as existent. Although we know that fictional characters do not exist, we cannot think of them as nonexistent, we have to think of them as existent.

> The idea of existence, then, is the very same with the idea of what we conceive to be existent. To reflect on any thing simply, and to reflect on it as existent, are nothing different from each other. That idea, when conjoin'd with the idea of any object, makes no addition to it. Whatever we conceive, we conceive to be existent. Any idea we please to form is the idea of a being; and the idea of a being is any idea we please to form. (Hume, *Treatise*, Book I, Part II, Section VI)

This characteristic of imagined objects counts in favor of the view that art is and can be only about the logically possible. Since impossible objects are necessarily nonexistent, we cannot imagine them as existent. And, if we cannot imagine them as existent, we cannot imagine them at all. Since imagination is constitutive for the art, art cannot be about the logically impossible.

Of course, nonexistent objects and people are too tempting theme to be missed. Italo Calvino wrote a novel *The Nonexistent Knight*. It is about *Agilulf*—the knight who does not exist. The point of the novel is that he is so virtuous and so perfect that he cannot exist. How can we conceive that? Well, there is an empty armor that talks, and that is him. In fact, *Agilulf* has some characteristics of the existent people and some characteristics of the nonexistent ones. This is how we can follow the story. Due to this fact, Calvino's *Agilulf* does not really violate Hume's insight that to imagine is to imagine as existent. After all, the novel is an allegory and the ontological pressure is not really hard. In the fantastic literature standards of possibility are more liberal. One might plausibly argue that the real content of the novel is not the nonexistent knight *Agilulf* but rather ideal stereotype of knight or something else.

6. *Dilemma and Options*

On the one hand, it seems that nothing can be about the logically impossible. While, on the other hand, it seems that art can be about anything, including logically impossible. So, the question is whether art can or cannot be about the logically impossible. In principle there are two main options here—positive and negative answer to the question. Call them *possibilism* and *impossibilism*.[12]

> *Possibilism* is the view that art is and has to be about the logically possible, it is not and it cannot be about the logically impossible.
> *Impossibilism* is the view that art does not have to be only about the logically possible, it can be about the logically impossible as well.

These two options are central claims of the views about the relationship between art and the impossible. Of course, they can be combined with other related insights and give a richer theories. I would point out five possible views about the issue. Notice: for the sake of brevity and style, in the rest of this article I will use "possible" primarily in the sense of "logically possible." Though, I believe that arguments and insights that hold for the logically impossible *mutatis mutandis* can be applied to the factually impossible.

1) One can accept possibilism and try to explain away the counterexamples. In this option, one has to show that that all the above mentioned examples are in fact possible.
2) One can accept possibilism and try to show that counterexamples are not really counterexamples but rather misdescriptions of possible situations. Since impossible cannot be represented, the impression of the impossibility must be a verbal matter or a matter of interpretation.
3) One can accept possibilism and argue that impossible situations should be seen as auxiliary artistic means for saying something about the possible. On this view, impossibilities could have their role in the art although the art would ultimately be about the possible.
4) One can weaken the requirement of possibilism and argue that the above mentioned examples, although impossible, are imaginable or *prima facie* conceivable. In this option, one has to argue that the subject matter of the art has to be imaginable or *prima facie* conceivable, although it does not have to be really possible.[13]

[12] In contemporary philosophical literature term *possibilism* denotes the view that all possible worlds are equally real, and it is opposed to *actualism*—the view that only the actual world is real. In spite of this terminological overlap, I have chosen this term because it perfectly expresses the idea.

[13] David Chalmers distinction between *prima facie* and *ideal* conceivability can be useful here: "S is prima facie conceivable for a subject when S is conceivable for

5) One can accept impossibilism and try to show that art can be about the possible just as it can be about the impossible. On this view, some states of affairs are possible, some are not, and art can be about both. Art can represent impossible situations.

Some of these views are compatible and can be combined into a single theory. 1) and 2) are compatible: some counterexamples to the possibilism can be explained away as in fact possible, and some as descriptions of something else. 3) is a certain concession to the impossibilism, although it is still in the general spirit of the possibilism: impossibilities are seen only as auxiliary artistic means. Although 4) is basically an impossibilist view, there is still a strong air of possibilism in it: although represented situation does not have to be possible, it has to be presented to us as possible and we have to experience it as possible. 5) is impossibilism in its simple and pure form.

The aim here is to provide (i) a single account that would hold for the whole of the art. Say, that 1) holds for all of the art. Though, the analysis might show that (ii) different accounts hold for different arts. Say, that 5) holds for the literature while 4) holds for the cinema and the theatre. Finally, it might turn out that (iii) there is no systematic account at all, and that relevant cases have to be analyzed one by one. As far as it goes, we will assume that there is a single account that holds for the whole of the art. If it turns out that there is no such thing, then one should withdraw to the less ambitious position

7. *Apparent and Real Content*

The interesting feature of the option 3) is that it relies on the assumption that work of art has the *apparent* and the *real* content.[14][15] There

that subject on first appearances. … S is ideally conceivable when S is conceivable on ideal rational reflection" (Chalmers 2002: 147). He also mentions van Cleve's distinction between *strong* and *weak* conceivability: "According to van Cleve, S is strongly conceivable for a subject when the subject sees that S is possible; and S is weakly conceivable when the subject does not see that S is impossible" (Chalmers 2002: 156).

[14] Sometimes it can be hard to tell what is work really about. A good example is Henry James' novella *The Turn of the Screw* from 1898. According to the supernaturalistic interpretation it is about the ghosts. According to the naturalistic interpretation it is about the main character's hallucinations. It is not clear which interpretation is right.

[15] Takashi Yagisawa believes that distinction between *semantics* and *pragmatics* might be useful here. Apparent content—actual meanings of words author uses would be semantics of the work, while real content—what author wanted to say would be pragmatics of the work. According to this view, to say that the leopard from the Hemingway's *Snows of Kilimanjaro* stands for human strivings or to say that characters from Bulgakov' *The Master and Margarita* stand for typical actual characters from the Soviet society in 1930's, is to do pragmatic analysis. Such approach might be appropriate and fruitful. Though personally I would rather link pragmatics to something more immediate and noninferential in the concrete circumstances of the utterance.

is a sense in which this assumption seems to be perfectly acceptable. Nobody thinks that *Aesop's fables* and Kafka's *Metamorphosis* are about zoology. Though, in some cases, this assumption might seem to be to interpretative. In *Terminator* fight between humans and machines would belong to its real content because it is possible, while time travel would belong to its apparent content because it is not possible. However, one might object that the time travel in *Terminator* is not an auxiliary means for achieving an additional artistic or commercial effect, but rather the constitutive part of the plot. Oxymoron is usually an auxiliary literary device and author who is using it is not ontologically committed to the existence of the corresponding impossible objects. However, it is not clear how authors whose central theme is time travel could eliminate time travel from the ontology of their works. If they write about the time travel then they write about the time travel. In the novel and the film *The Time Traveler's Wife* random time travels of the main character determine the plot structure. They are constitutive for the work. Nevertheless, one might argue that we understand the plot *by analogy* with the actual people who, due to the nature of their jobs, do not spend time with their families. Sailors, soldiers, construction workers, travelling salesmen, prisoners, emigrants, … they are all absent from their homes for months. We understand what time traveler's wife goes through by analogy with what wives of these people go through. One could replace random time traveler with the CIA agent who in the middle of the night receives a phone call and suddenly has to leave on an undetermined period of time. The plot structure would basically remain the same. Time travel is just an additional spice to the story. The same holds for the *Terminator*. It does not matter whether assassins come from different time or from different town. Plot structure remains the same.[16] At this place, one can accept 2) and argue that random time travel is just a misdescription of usual father's absence. Or, one can accept 3) and argue that time travel is a legitimate artistic auxiliary means which belongs to the apparent content and therefore has no importance in the ontological analysis of the film. These two claims are very close and the difference might even be purely verbal in the end. If we cannot understand an impossible situation, the question is what is really going on when we seemingly understand it. According to 2), since *we cannot understand* an impossible situation, we understand something else—some similar but possible situation. According to 3) *we understand* an impossible situation in the sense that it tells us something about the possible.

All in all, the consequence of 3) is that the requirement that fictional character and plots could have existed does not apply to the apparent content of work of art, but only to the real content.

[16] Perhaps *The Time Traveler's Wife* and *Terminator* can be plausibly explained away, but in some cases time travel seems really constitutive for the plot: BBC TV series *Doctor Who*, or Zemeckis' trilogy *Back to the Future*, not to mention H.G. Wells' novella *The Time Machine*. Nevertheless, if time travel is not logically possible, they have to be about something ese.

> A fictional character or plot has to be possible only if it belongs to the real content of the work of art.

For this reason the claim does not apply to Aesop's talking fox and Kafka's transforming man. They could not have existed but this does not matter because they are only auxiliary artistic means for saying something else and do not belong to the real content of the work. Cartoon *Madagascar* is not about the animals. It is about the people. This is the way to preserve possibilism about the art on the one hand, and combinatorial or some other realistic theory of possibility on the other hand. This way may be demanding and hardworking because it requires deep and detailed analysis of works of art, but for the same reason it may be fruitful and rewarding because it will deepen our understanding of the works of art.

8. *Art and the Impossible*

8.1. *Can Jean-Luc Picard fight himself?*

In science fiction TV series *Star Trek: The Next Generation* captain of the USS Enterprise Jean-Luc Picard often encounters himself. He encounters his past and future selves, or his possible selves.[17] Given his strong-willed character, these encounters can end in a fight. The question is what are we watching when we are watching such scenes. Options 1) and 4) do not have much chance because we can clearly see two men on the screen, and the assumption is that there is only one. The contradiction is obvious, and the situation cannot be possible, nor *prima facie* conceivable. Though, we are given some explanation of the situation: "due to the fluctuations in the force field, blah, blah, blah." So we might have the impression that the situation is possible. But since the end result is a contradiction, processes that bring it about cannot be possible.[18] It seems that in such situations 2) might have a chance. A scene on the screen can be described as a scene in which Jean-Luc Picard from time t_1 fights Jean-Luc Picard from t_2, or a scene in which Jean-Luc Picard from world w_0 fights Jean-Luc Picard from w_5. But that must be a *misdescription* of what we really see on the screen. What we

[17] Conceptual artist Dalibor Martinis interviewed his future self. He recorded the questions in 1978 and gave the answers in 2010 (https://www.youtube.com/watch?v=qWeVIcLE0Kg). Perhaps this interview can be verbally described in a paradoxical manner, as a conversation between earlier and later self, but that would be a misdescription of what was really going on.

[18] A proponent of 1) could rely on four-dimensionalist ontology and argue that what we have at the screen are two *temporal slices* of Picard. Since temporal slices have existence of their own, independently of the temporal wholes to which they belong, it is in principle possible that they encounter. Four-dimensionalism is an overall better background ontology for the time travel cases because it stipulates that past, present, and future are equally real. It suits well for the time travel cases but not for the travel between different possible worlds, which is considered as impossible. Although four-dimensionalism is very popular view these days, I will not discuss it here.

really see on the screen are *two guys* that look exactly the same and fight each other, not *one and the same guy* fighting himself. Both guys are in the same time and in the same possible world. It is not one and the same guy from different times and different possible worlds. However, the problem with this option is that it does not take into account the initial assumption that two men on the screen are supposed to be one and the same man. For this reason it cannot be taken as a good reconstruction of what we see at the screen. According to 3) situation on the screen might show the differences in values and aims of the younger and the older Picard. Younger one would take an adventurous course of action, while the older one would be more cautious. Since they are both stubborn, they fight. So, although the situation on the screen is impossible, it represents something that is possible. The fight on the screen is in fact an inner fight. However, the problem with this interpretation is that the fight is not inner. These episodes are not about Picard's dilemmas or inner struggles. They are about objectively existing people and their actions, in the fiction of course. 5) might seem like the most promising option here. What we have on the screen is an impossible situation but it does not matter. We watch it, we enjoy it. Who cares whether it is possible or not. This position might seem like an easy way out but what counts against it is its overall implausibility. How can we watch something that is not possible? Can we watch a round square?

However, if we take a look at the plots of such episodes, we will see that they are usually about *the attempts to take over command over the ship*. As far as the plot is concerned, it does not really matter *who* tries to take over the ship. Is it a Star Fleet colleague, a younger officer, a pirate, a member of unknown species, … or Picard himself (younger, older, or possible). What really matters is that *somebody* tries to take over the ship and that (our) Picard fights against it. The fact that the one who tries to take over the ship is Picard himself (younger, older, or possible) is an extra ingredient that makes no real difference to the plot. This extra ingredient is dispensable and therefore does not belong to the real content of the episode, but only to its apparent content.

> If the assumption that A and B are identical can be given up without affecting the plot structure, then it does not belong to its real content.

This analysis is an instance of the option 3). Though, the real content is not Picard's inner struggle but his willingness to keep command over the ship. This shows that we can be wrong in the analysis of the real content. In accordance with 3) we might conclude that Jean-Luc Picard fighting himself is an impossible situation which represents something possible—his determination to keep control over his ship.

In this context one more thing should be mentioned. The assumption of the argument is that *our* Jean-Luc Picard from our time and our dimension *is identical* to the one who comes from other time or other dimension. However, this is not the impression that we have when we

watch it. As spectators we are emotionally attached to *our* Jean-Luc Picard, not to other guys who come from different times or dimensions. We regard them as *copies*, as *aliens*, not as men who are identical to the *original*. We want *our* guy to win and to throw them back to where they came from. Although two guys are supposed to be identical, our reactive attitudes toward them are not the same. Perhaps the assumption of identity should be revised here.

8.2. *Can Anna Karenina throw and not throw herself under the train?*

Say that there is a novel in which Anna Karenina does and does not throw herself under the train. In Chapter 8 she comes to the railroad station and throws herself under the train. In Chapter 9 she comes to the railroad station, changes her mind, and goes away. On its face value this is a direct contradiction and no account of possibility can render it possible, no matter how liberal the account is. Contradiction is so obvious that it cannot be *prima facie* conceivable. So, options 1) and 4) drop off. It is not clear how this could be a misdescription of something else. Since this is a novel, there can be no discrepancy between the text and the visual content. Inconsistency is so obvious that it is not possible that the author overlooked it or that he misexpressed himself. So, 2) drops off as well. It might seem that the proponent of 5) has his chance here. He might argue that the novel *is* about a woman who did and did not throw herself under the train. But since such a woman cannot exist, how could a novel be about *her*? There can be no *her* here. Since no women can satisfy both descriptions, a novel must be about *nobody*. Impossibilist would accept that *she* is not possible, but would argue that nevertheless the novel is about *her*—a woman who cannot exist. Generally, the impossibilist believes that *there are* impossible objects. They cannot exist but we can think and talk about them. So, the novel *is* about a woman who at the same time does and does not throw herself under the train. Possibilist would argue that necessarily non-existent subject matter cannot be a subject matter at all. Therefore, such a novel is either about *nothing* or about *something else*. Since it cannot be about nothing *it has to be about something else*! Do we have a metaphilosophical stalemate here? Possibilist and impossibilist would insist on their claims and accuse each other of begging the question. P: She cannot exist and the book cannot be about her. I: She cannot exist and the book is about her! … Although a situation might seem symmetrical here, I think we should reject impossibilism on the grounds of its overall implausibility. So, the only option left is 3) and it seems that it fits the bill. According to this option, a novel is not about contradictory state of affairs, it is about something else. Perhaps author wants to express her dilemma—to live or not to live. Perhaps author wants to emphasize the fact that one always has a choice. Maybe author wanted to describe *forking paths* of her life—possible courses of action that she could have

taken. One might argue that the novel expresses many world hypothesis or some other recent insight of quantum mechanics, but I would not go that far. Anyway, the point is that an impossible situation is used as artistic means for saying something about the actual or the possible.

We mentioned author's intentions here. Impossibilist might argue that if author's intention was to describe something impossible, we have to admit that the work of art is about something impossible. Since author's intentions determine the content of the work, if the author wants to describe an impossible event, then his work describes an impossible event. However, even if the author wants to describe something impossible, the question is whether he can be successfully in it. A writer can produce inconsistent text, that's for sure. But the question is whether that text would stand for something. Is there anything that such a text would describe? No! So, even if the author's intention really was to describe an impossible event, the text that he produces could not be a description of an impossible event. He can try, but he cannot do it. He would necessarily be unsuccessful. If somebody wants to draw a round square, good luck to him! In 2002 Roy Sorensen offered a 100$ prize "to the first person who identifies a picture of a logical impossibility." (Sorensen 2002: 337) To my knowledge the prize is still unclaimed.

8.3. *Can water fall in a circle?*

One might say that works of Dutch graphic artist M. C. Escher represent impossible objects. On his lithographs *Waterfall*, *Relativity*, *Ascending and Descending*, water flows uphill, people endlessly climb in a circle, gravity works in different directions, perspectives are messed up, etc. If his works really represent impossible objects, then art can be about the impossible and impossibilist wins this debate. So, the question is whether his works really represent impossible objects.

It seems that strategy 1) cannot work in Escher's case. The content of his graphics can hardly be explained away as possible, even under the assumption of maximally liberal criterion of possibility. Perhaps works of some similar authors might be explained away as logically possible. One might argue that it is logically possible that giant rock on Rene Magritte's *Castle in the Pyrenees* just floats in the air, unsupported by anything. This is certainly physically impossible, perhaps metaphysically impossible, but logically possible. Since a world without the gravity is not contradictory, *Castle in the Pyrenees* is logically possible. We can imagine it, we can conceive it. However, in the Escher's case things seem to be different. The content of his graphics can be described as contradictory: water that floats in a circle creates a waterfall, it floats upward and downward at the same time, people that climb up the stairs do not climb up the stairs, etc. These descriptions are logically contradictory, and they cannot be explained away by appeal to liberal criterion of possibility.[19]

[19] It is very interesting question whether pictures are discursive. Sorensen says: "I agree with most philosophers in denying that pictures are discursive" (Sorensen

Does it mean that Escher was successful in *drawing* or *representing* impossible objects? Such a conclusion would be very hard to accept because graphics is visual and we cannot visualize impossible objects.

> If drawing X demonstrates the possibility of X, then we appear to have a quick proof that it is impossible to draw an impossible object. Drawing an impossible object would show that it is possible for an impossible thing to exist. Contradiction. Therefore, it is impossible to draw an impossible object. (Sorensen 2002: 343)

In accordance with 2) one might argue that these contradictory descriptions are in fact misdescriptions of something else. Since graphics as physical objects cannot be contradictory, the seat of contradictions must be in our interpretation, not in objects themselves. There cannot be anything contradictory in the inked and non-inked areas of the lithography. However, the problem with this line of reasoning is that we do not discuss Escher's graphics as physical objects, we discuss their content, we discuss what they represent. The question is whether they *represent* something contradictory, not whether they as physical objects are something contradictory.

A natural thing to say is that Escher's graphics do not represent impossible objects but rather create *illusions* of impossible objects. This amounts to saying that these works only *seem* to represent impossible objects but in fact they do not. But, if these works do not represent impossible objects, what do they represent? What are they really about? It is not sufficient to say that they are about something else, we have to say what are they about. There are several answers possibilist might give here. (1) One might argue that *Waterfall* represents waterfall, that *Ascending and Descending* represent monks walking up and down the stairs, etc. Here one would simply ignore the perspectival inconsistencies and say what is the object that graphics represent. In the same way in which we ignore minor ungrammaticalities in the language. Of course, perspectival inconsistencies and paradoxicality was Escher's point and we cannot ignore it. (2) One might say that Escher's graphics are about possible elements composed in a wrong way. Though, it is not clear whether possibilist is allowed to give such answer. If the wrong way means the way that brings about impossible objects, possibilist cannot give such answer. If the wrong way means that author uses more than one perspective instead of only one, this path seems to be open to the possibilist. This path would lead us to the complete and detailed description of the graphics without any mention of inconsistency and paradoxicality. But again, such a description would miss author's intention. (3) One can argue that Escher's graphics are really about the role of perspective in the perception. For they show what happens if the rules of perspective are not obeyed. Perceptual illusions are fun and they can be used for amusement only. But their real purpose is

2002: 341). However, the claim here is weaker. The claim is not that pictures are discursive in their nature, but that their content *can be described* with sentences.

to reveal the mechanisms of our perception: if we see where things go wrong, we can guess how the underlying mechanisms work when everything is right. And this is the true role of illusions in the psychology of perception. In the same way, the true role of Escher's graphics is to make us aware of the role and nature of perspective in the perception. This is what Escher's graphics are really about. There is certainly some truth in this insight but what remains to be seen is whether Escher's graphics do so by representing impossible objects. (4) A similar answer would be to claim that Escher's graphics represent Penrose's stairs and Penrose's triangle. This is what they are really about. In the *Waterfall* there are three Penrose's triangles one above the other. This is obviously true but, again, the question that is relevant in this discussion is whether Penrose's stairs and Penrose's triangle are impossible objects. A proponent of 3) would allow that. He would argue that Penrose's stairs and Penrose's triangle are impossible objects that serve as artistic means for saying something about the actual. But here we are still dealing with 2), which is purely possibilist option. (5) One can argue that Escher's graphics express ideas. *Relativity* represents relativity—the idea that things can be seen from more than one perspective. In this sense, Kurosawa's *Rashomon* and Escher's *Relativity* express the same idea. Escher himself took *Ascending and Descending* to express life—endless motion in a closed circle. One might see *Ascending and Descending* as representing *Wheel of Fortune* or *Rota Fortunae*. Though, since there are no desired or undesired objects in a circle, it should rather be taken to depict existential meaninglessness. (6) Finally, one might argue that graphics do not represent anything. There are two independent reasons for saying this. First, since objects that graphics purport to represent necessarily cannot exist, graphics represent *nothing*. Second, objects should be drawn from a single perspective. If they are not, they cannot represent anything. Just like ungrammatical sentences say nothing, graphics drawn from more than one perspective represent nothing. In spite of its philosophical elegance, this view is hard to defend. How can we say that they represent nothing when obviously there is something on them? We can recognize some forms and shapes. It is something, it cannot be nothing. Anyway, a proponent of 2) has to argue in favor of some of these, or some similar answers. Since he claims that Escher's graphics cannot be about the impossible objects, he has to say what are they really about.

A question that is relevant here is what does it mean that graphics G represents object X. Representing can have two meanings here, transitive and intransitive. To say that G represents X in the transitive sense implies that outside of G and independently of G there exists an object X. To say that G represents X in the intransitive sense implies that X's existence is within G, not independent of it. Nelson Goodman says that "the picture of" and "represents" can be understood as one place predicates and as two places predicates.

> What, for example, do pictures of Pickwick or of a unicorn represent? They do not represent anything; they are representations with null denotation. Yet how can we say that a picture represents Pickwick, or a unicorn, and also say that it does not represent anything. … Obviously a picture cannot, barring equivocation, both represent Pickwick and represent nothing. But a picture may be of a certain kind—be a Pickwick-picture or a man-picture—without representing anything. (Goodman 1968: 21–22)

Escher's graphics obviously cannot represent impossible objects in the transitive sense of represent, because these objects are not out there to be represented. The question is whether Escher's graphics represent impossible objects in the intransitive sense of the word.

In the case of Escher's graphics there is one more ambiguity to be clarified. The question whether we can *see* the impossible objects on these graphics can be understood in two senses.

1) We can see that such object are not possible.
2) We can see impossible object.

Possibilist is satisfied with 1). By looking at the graphics we can see what the author purports to represent and we can see that such object cannot exist. Impossibilist needs sense 2). He has to claim that artist did succeed in representing an impossible object and that we can see it on the graphics. Now, the crucial question is whether we can *see* an impossible object on the graphics or not. On Escher's graphics, elements are represented realistically but then combined in impossible ways. Parts are possible while wholes are not. In the top center of the *Relativity* there are two guys walking on the same stairs, in the same direction, but one climbs up the stairs while another one goes down the stairs. Each of these guys taken in itself is possible, but if they are taken together they are not possible. We can *see* one of them going down the stairs. We can *see* another one going up the stairs. But can we *see* both of them at once, going up and down the stairs at once? Impossibilist should argue that we *can* see them both at the same time, while possibilist should argue that we *cannot*. For if we cannot see them, then graphics does not represent them. Talking about the *Relativity* Hofstadter says that we do see the impossible combinations.

> So we are forced, by the hierarchical nature of our perceptive processes, to see either a crazy world or just a bunch of pointless lines. A similar analysis could be made of dozens of Escher picture, which rely heavily upon the recognition of certain basic forms, which are put together in nonstandard ways; and by the time the observer sees the paradox on a high level, it is to late—he can't go back and change his mind about how to interpret the lower-level objects. (Hofstadter 1979: 98)

However, it seems that Hofstadter goes wrong here. In the case of duck-rabbit and other *Gestalt* illusions we *either* see the duck *or* we see the rabbit. We cannot see both at the same time. Possibilist will argue that the same hold in the case of *Relativity* and other Escher's graphics. We *either* see a man going down the stairs *or* we see a man going up the stairs. We cannot see both of them in the same time. Therefore, we can-

not see an impossible situation. And if we cannot see it, the graphics cannot represent it. Seeing includes subconscious inferential processing and, although pictures of both men are reflected on our retinas, we cannot form a single picture of both of them. This is not a matter of contingent limitation of our perceptual apparatus, it is a matter of necessity. For, to see X as a single object just is to see that its elements exist together. And if we cannot see that its elements exist together, we cannot see it at all. In one moment we see one element of the purported impossible object, in another moment we see the other element, but there is *no moment* in which we clearly see the impossible object. We can sum up this insight in the following principle.

If we can see that X is impossible, then we cannot see X.

This may sound like a contradiction, so here we have to clarify the sense in which we can see that something is impossible. When we say that we can see the elements, we talk about seeing in the direct and *perceptual* sense. However, when we say that we can see that elements cannot coexist, we talk in the indirect, *intellectual* or *inferential* sense of seeing. Strictly speaking, the impossibility of coexistence of the elements is not something that we can literally see, but rather something that we *infer*. Since this inferential process is mainly subconscious and perception related, we can say that we can *see* that elements cannot coexist, but this is seeing in its secondary and derivative sense.

This analysis shows that impossibilist does not have a case with Escher's graphics. Escher certainly created very interesting and nice puzzles, but he did not represent impossible objects. We can say that he was very good in *trying* to draw impossible objects, but not in actually drawing them. He did get as close as possible to representing impossible objects, but he did not really represent them. So, in the case of Escher's graphics the most promising option seems to be 2).

8.4. *Can we kill our own grandfathers?*

The paradox of patricide is standard *a priori* argument against the possibility of time travel. It is an *reductio* argument: if it was possible to travel in time, then it would be possible to travel back in time and kill one's own grandfather before he conceived one's own father. But this is absurd: if one kills one's own grandfather before he conceived one's own father, one could not exist and therefore one could not kill anybody. Since this consequence is absurd, time travel is impossible, so runs the argument. Nevertheless, we can easily imagine a novel or a film with exactly such a plot. Guy finds out that his grandfather was very abusive man who was heavily beating grandmother during years of their marriage. Our guy gets very angry on his grandfather and, since time travel is technologically possible, decides to prevent past grandmother's suffering. He gets a gun, jumps back in time before his grandfather met his grandmother, finds his grandfather and kills him.

After that he gets back into the present and normally continues his life. Everything is the same as it was, only his grandmother sometimes has unexplainable feeling of gratitude toward him.

Impossibilist 5) would argue that this case clearly shows that art can be about the impossible. Plot is impossible but nevertheless we can watch the movie, approve or disapprove of the protagonist's decision, feel the suspense and the relief, etc. What difference does it make whether the plot is possible or not? This is an impossible situation and the film *represents* it. It is *about* an impossible course of events!

At this place possibilist might argue that the film represents *elements that are by themselves possible*: it is possible that one feels sorry for one's grandmother, it is possible that one kills an abusive man, etc. The film is about these elements, not about the impossible way in which these elements are supposed to be combined. According to this view – 2), to say that the film is about the paradox of patricide would be to misdescribe the real content of the film. The film is about the guy who kills an abusive man, not about the guy who kills his own abusive grandfather before he met grandmother. Of course, impossibilist 5) would protest here and argue that we should not reinterpret the content of the film. If the screen writer says that guy kills his own grandfather, if the director says that guy kills his own grandfather, if the narrator in the film says that guy kills his own grandfather, … then guy kills his own grandfather. A philosopher should not patronize all these people and tell them what they really said. They said what they said! However, it seems that possibilist 2) has a strong point here: if *X* is impossible, one cannot represent it. Just as painter cannot represent a round square, writer or movie maker cannot represent a paradox of patricide. What they really represented has to be something else. They can *think* that they represented a paradox, they can *say* it, but they cannot do it. Just like Escher in his graphics, they can represent elements of an impossible situation but they cannot represent an impossible situation.

Possibilist may take course 1) and try to show that the paradox of patricide is not really a paradox. One might claim that the universe contains some sort of buffering mechanisms that would fix the disturbances we may cause when we travel in time, or something like that. But, if it is a matter of necessity that we have parents that we have, it is not clear what these mechanisms should do. With different grandfathers our parents would not be our parents and we would not be we. So, option 1) does not seem promising in this case.

4) might seem promising here. One might argue that the content of the film *prima facie* seems consistent: when we watch the film and enjoy it, we simply do not think about its paradoxical consequences. Plot of the film *seems* consistent to us although it is not. We only *think* that we watch a guy who kills his own abusive grandfather, but we do not. What we really watch is something similar—a guy who kills an abusive man, not a guy who kills his own abusive grandfather. But if this is so,

then the film cannot be *about* a guy who kills his own abusive grandfather. It cannot *represent* such a guy. And this shows that position 4) cannot stand on its own and that it falls back into the possibilism.

One might take stance 3) and argue that in this case the paradox of patricide is just a means for saying something else. Perhaps for saying that we cannot wash away shame for the sins of our ancestors, that there exists tragic guilt, moral luck, original sin, or something like that. One might argue that the very impossibility of the plot shows that our protagonist will forever remain the grandson of a molester and that there is nothing he can do to change it. The message of the film would be that we have to live with the sins of our fathers! Of course, whether this is the message of the film would significantly depend upon the intentions of the author. So, the paradox of patricide can be only an additional spice to the otherwise consistent story of killing a molester, or it can be a means for saying something important about the human condition. That will depend upon the intentions of the author and the reception of the audience.

The interesting question here is whether 2) and 3) are compatible or not. If author cannot represent an impossible situation, how can he use it to say something about the possible? He can use it only if he can represent it. Does it mean that 3) is an essentially impossibilistic option? Appeal to the distinction between apparent and real content can be of no use here. Even if an impossible situation belongs to the apparent content, it is still not clear how can author use it if he cannot represent it? However, possibilist might try to argue that author can use an impossible situation without representing it. In that case the claim would be that author represents elements of an impossible situation, not an impossible situation itself, and that we *infer* that certain combination of elements is impossible. Speaking in Wittgensteinian way, one might say that in such cases author *indicates* an impossible situation without representing it. Here we have a difference between (1) showing that an object is not possible and (2) showing an impossible object. (1) is compatible with possibilism while (2) is not. The point is that (1) is sufficient for the analysis of a film or a novel with such content.

To sum up this paragraph, it seems that a work of art that is *prima facie* about the paradox of patricide cannot really be about it. Hence, it cannot support impossibilism – option 5). Possibilism, options 2) and/or 3), provide better reconstruction of such novel or film.

9. *Can we individuate different impossible states of affaires?*

If inconsistent stories do not represent anything, then we cannot individuate them according to their content because they all have the same content—nothing. However, on the other hand, it seems that different inconsistent stories do have different contents and that we can individuate them. A novel about Anna Karenina who does and does not throw

herself under the train, and a novel about Captain Ahab who does and does not catch Moby Dick, although both inconsistent, have different contents. The first one is about Anna Karenina and her suicide, and the second one is about Captain Ahab's catching a big white whale. Story about a guy who travels back in time and kills his grandfather before he met grandmother is about a guy who travels back in time and kills his grandfather, and so on. Prima facie, these stories obviously have different contents and we can individuate them. Impossibilist might try to build his case here. He might try to argue that different stories are different stories, no matter whether they are about the possible courses of events or about the impossible ones. On the impossibilist's view, the impossible stories are just as good as the possible ones. Therefore, art can be about the impossible too.

This view may have some *prima facie* plausibility but it is ultimately untenable. Inconsistent descriptions have no content and we cannot individuate them by their content. As we saw, impossible situations are composed of possible elements arranged in an impossible way.[20] We individuate impossible situations by appeal to the *possible elements they are composed of*, not by appeal to different impossible situations they make part of. There are no different impossible situations. There are only different possible elements that we can try to arrange in an impossible way, try but not succeed. We can individuate Anna Karenina that throws herself under the train, we can individuate Anna Karenina that does not throw herself under the train, but we cannot individuate Anna Karenina that does and does not throw herself under the train. Such Anna cannot exist and we cannot think and talk about her. Talk about such Anna cannot be about her because there is no and there can be no her. It cannot be about anything. The following might sound strange but it is so. Say that Anna Karenina is written in two volumes. In volume one she throws herself under the train and in volume two she does not. In this case volume one would have content, volume two would have content, but both volumes taken together would have no content at all. Impossibilist might protest here: How could that be? How could parts be meaningful and whole meaningless? The answer is simple. If volume two negates what volume one says, then both volumes taken together just say nothing. For this reason we cannot individuate different impossible situations.[21] Can we tell apart Anna who does and does not throw herself under the train from Anna who does and does not marry Vronsky?[22] Well, loosely speaking, we might say that the first one is

[20] Raymond Smullyan asks: "What happens if an irresistible cannonball hits an immovable post?" A contradiction happens! An irresistible cannonball can exist, an immovable post can exist, but they cannot both exist at the same time (Smullyan 1978: 8).

[21] In other words, there cannot be different impossible worlds. In a charitable interpretation we could individuate them, but only by appeal to their possible ingredients, not by appeal to their impossible composition.

[22] Quine was worried that we cannot individuate possible objects (Quine, 1948). Perhaps so, but we certainly cannot individuate impossible ones.

about suicide while the second one is about marriage. Corresponding novels would contain some textual difference and we could individuate different texts. However, strictly speaking, these two novels would not be about different characters, they would not be about characters at all. Therefore, impossibilist cannot build his case on these grounds. Art cannot be about the impossible. It has to be about the possible.

10. *Conclusion*

I hope that I have successfully shown that art is not and cannot be about the logically impossible. Since art is typically about something, and logically impossible is nothing, works of art that are *prima facie* about the logically impossible have to really be about something else. People who claim that art is and can be about the logically impossible are going too far. Art cannot depict logically impossible events, plots, objects, or worlds because they are necessarily nonexistent and nothing can depict them. Though, logically impossible can be used as auxiliary means that points to something possible.

References

Alber, J. 2016. *Unnatural Narrative: Impossible Worlds in Fiction and Drama*. University of Nebraska Press.

Aristotle 2015. *Poetics*. Translated by S. H. Butcher. Calgary: Scholar's Choice.

Armstrong, D. M. 1989. *A Combinatorial Theory of Possibility*. Cambridge: Cambridge University Press.

Chalmers, D. 2002. "Does Conceivability Entail Possibility?" In T. S. Gendler and J. Hawthorne (eds.). *Conceivability and Possibility*. Oxford: Oxford University Press.

Goodman, N. 1968. *Languages of Art*. Indianapolis and New York: The Bobbs-Merrill Company.

Hofstadter, D. R. 2000. *Gödel, Escher, Bach: and Eternal Golden Braid*. New York: Penguin Books.

Hume, D. 1985. *A Treatise of Human Nature*. New York: Penguin Classics.

Kripke, S. 1980. *Naming and Necessity*. Oxford: Blackwell.

Lewis, K. D. 1986. *On the Plurality of Worlds*. Oxford: Blackwell.

Loux, M. (ed.). 1979. *The Possible and the Actual*. Ithaca: Cornell University.

Mellor, D. H. 1998. *Real Time II*. New York: Routledge.

Mondadori, F. and Morton, A. 1979. "Modal Realism: The Poisoned Pawn." *Philosophical Review* 85 (1): 3–20.

Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford university Press.

Plantinga, A. 1978. *The Nature of Necessity*. Oxford: Oxford University Press.

Priest, G. 1997. "Sylvan's Box: a Short Story and Ten Morals," *Notre Dame Journal of Formal Logic* 38 (4): 573–582.

Quine, W. V. O. 1953. "On What There Is." In *From a Logical Point of View*. Cambridge: Harvard University Press.

Sider, T. 2009. "Time." In S. Schneider (ed.). *Science Fiction and Philosophy*. Oxford: Wiley-Blackwell.

Smullyan, R. 1978. *What is the Name of This Book?* Hoboken: Prentice-Hall.

Sorensen, R. 2002. "The Art of the Impossible." In T. S. Gendler and J. Hawthorne (eds.). *Conceivability and Possibility*. Oxford: Oxford University Press.

Thomasson, A. 1999. *Fiction and Metapysics*. Cambridge: Cambridge University Press.

Croatian Journal of Philosophy
Vol. XXI, No. 61, 2021
https://doi.org/10.52685/cjp.21.1.10
Received: November 23, 2020
Accepted: March 15, 2021

Generative Linguistics Meets Normative Inferentialism: Part 2

DAVID PEREPLYOTCHIK*
Kent State University, Ohio, USA

This is the second installment of a two-part essay. Limitations of space prevented the publication of the full essay in a previous issue of the Journal (Pereplyotchik 2020). My overall goal is to outline a strategy for integrating generative linguistics with a broadly pragmatist approach to meaning and communication. Two immensely useful guides in this venture are Robert Brandom and Paul Pietroski. Squarely in the Chomskyan tradition, Pietroski's recent book, Conjoining Meanings*, offers an approach to natural-language semantics that rejects foundational assumptions widely held amongst philosophers and linguists. In particular, he argues against extensionalism—the view that meanings are (or determine) truth and satisfaction conditions. Having arrived at the same conclusion by way of Brandom's deflationist account of truth and reference, I'll argue that both theorists have important contributions to make to a broader anti-extensionalist approach to language. Part 1 of the essay was largely exegetical, laying out what I see as the core aspects of Brandom's normative inferentialism (1) and Pietroski's naturalistic semantics (2). Now, in Part 2, I argue that there are many convergen-*

* My deepest thanks to Daniel Harris, Ryan DeChant, Elmar Unnsteinsson, Eliot Michaelson, and Jessica Lynne Cooperrider for helpful advice on earlier drafts of this paper; thanks also to Nate Charlow, Matt Moss, and other participants of the 2018 Ontario Meaning Workshop, for stimulating exchanges on related topics. That said, if some of what I claim here is false, then these people are all partly to blame. Ditto for cases where I say something awkwarder than usual; they should of caught that. My sincere gratitude, lastly, to all of the participants at the two symposia at which I was fortunate to present earlier versions of these ideas, including the 2018 Croatian Meaning Workshop. From the latter group, I owe special thanks to Paul Pietroski (obviously), Michael Devitt, John Collins, Anna Drożdżowicz, and, most of all, to Dunja Jutronić—not only for inviting me to participate in the workshop, but for her saintly and near-inexhaustible patience with my obnoxious delays in submitting the present work to the Journal. Not to mention the outrageous word-count. I beg the reader's forgiveness for the latter, as well.

ces between these two theoretical frameworks and, contrary to first appearances, very few points of substantive disagreement between them. If the integration strategy that I propose is correct, then what appear to be sharply contrasting commitments are better seen as interrelated verbal differences that come down to different—but complementary—explanatory goals. The residual disputes are, however, stubborn. I end by discussing how to square Pietroski's commitment to predicativism with Brandom's argument that a predicativist language is in principle incapable of expressing ordinary conditionals.

Keywords: Generative linguistics; anti-extensionalism; normativity; inferentialism; predicativism; public language; communication.

3. *Prospects for Ecumenicism*

Introduction

In Part 1 of this essay, we surveyed the core commitments of two large-scale theoretical frameworks in the philosophy of language—those of Robert Brandom and Paul Pietroski—and seen some of the ways in which they play out in the realm of semantics, including in detailed analyses of various linguistic constructions. It may appear that the two views are so different in substance and overall methodology that a conversation between the two is unlikely to bear much fruit. This is a large part of why so few conversations of this kind ever take place. In the present section, I argue for the contrary perspective, outlining an ecumenical approach that seeks to integrate the two in a variety of ways. In surveying what I take to be significant points of convergence—which then serve as background for constraining residual disputes—I rebuff various superficial objections to the possibility of integration. In each case, I show how the theoretical differences that they point to can be reconciled without doing much (if any) violence to either view.

3.1 *Truth, reference, and other non-explanatory notions*

One obvious shared commitment between Brandom and Pietroski—indeed, the one that most clearly motivates the present enterprise—is their common rejection truth-conditional semantics. Let's review the key points and add some new ones.

Pietroski surveys a battery of arguments against Davidson's proposal, including its more recent incarnations in possible-worlds semantics. These include troubles with (i) empty names, (ii) co-extensive but non-synonymous expressions, (iii) polysemy, (iv) compositionality, (v) liar sentences, and (vi) event descriptions (*inter alia*). Brandom's skepticism is more foundational. On his view, an explanatory strategy that takes truth and reference—conceived of as "naturalizable" word-world relations—as fundamental semantic notions will require a metase-

mantics that is, at best, optional, at worst, incoherent, and, at present, non-existent. Although Brandom doesn't pretend to have supplied a knock-down argument against it, the flaws he identifies in the various attempts to work out this strategy strike me as fatal. Coupled with his development of a powerful alternative—a large-scale framework constructed from the top down, with pragmatics taking an unconventional leading role—as well as his well-motivated treatment of the notions of truth and reference, Brandom deals a serious blow to the mainstream approach of subordinating pragmatics to semantics.

Brandom's *main* reason for pursuing a normative metasemantics is the inability of a purely descriptive ("naturalistic") account to capture the normative notion of (in)correct rule-following. But this is not his *only* argument, and it's worth taking a moment to spell out what strikes me as a potentially more powerful consideration—one that, in making fewer assumptions, can appeal to theorists of a broader stripe.

All extant attempts at "naturalizing" meaning, content, representation, and the like, have in common their insistence on employing in only *alethic* modal notions in their analyses. These include dispositions (Quine), causation (Stampe), causal covariance (Locke), natural law (Dretske), nomic possibility (e.g., Fodor's "asymmetric dependence"), and even appeals to non-normative teleology (Cummins and, independently, Millikan). Brandom points out that, even if these *could* account for ordinary empirical concepts, it's not at all clear how they might be extended to the very concepts that appear in the analyses—i.e., the *alethic modal* notions just listed (among others).

While it's possible to envision, if only dimly, how something in a person (or a brain) can causally co-vary with—or bear nomological relations to—water, mountains, and even crumpled shirts, there's simply no naturalistic model for envisioning the relation between, on the one hand, the words 'possible', 'disposition', or 'asymmetric dependence', and, on the other hand, any particular set of things, events, or phenomena out in the natural world. The same is arguably true for logical, mathematical, semantic, and deontic vocabulary. (Recall the fonts, functions, and fears from 2.3.) Indeed, the metasemantics doesn't seem to get us much farther than 'stick' and 'stone'. And the experts seem to have given up, since the late 1990s, on the hard at work of bringing 'fail' and 'decisively' into the fold. One is well-advised not to bank on any striking new developments in this area, unless, of course, something dramatic happens in the surrounding domains of inquiry. (My money, for what it's worth, is on the AI people.)

By contrast, a metasemantics that makes explicit and essential use of *normative* terms—paradigmatically, deontic modals—is ultimately able to "eat its own tail", to use Brandom's imagery, by shoring up a principled account of *those very notions*. As discussed in 1, Brandom treats normative expressions as serving the function, characteristic of logical vocabulary more generally, of expressing (i.e., making explicit) one's commitments to the propriety of an inference or a plan of practi-

cal action. Brandom (2008) offers, in addition, an account of *alethic* modal vocabulary (recall the safety measures for gasoline wicks in 1.2), as well as a detailed formal analysis of its many important relations, both semantic and pragmatic, to the deontic variety. In this way, his account can claim a major advantage over virtually any *conceivable* attempt at a naturalistic alternative. And, again, the force of this argument does *not* depend on a prior assumption about the normativity of meaning. This functions here not as a premise, but as a conclusion.

3.2 *Naturalism*

Residual worries about adopting a normative metasemantics will doubtless trouble self-avowed naturalists (including former versions of the present author), who tend to have a constitutional aversion to trafficking in normative considerations. But this, too, should be tempered—or so I'll now argue. The concern is, to my mind, largely dampened by the fact that Brandom's norms are in no way "spooky" (despite drawing heavily on Kant), but, rather, grounded directly in social practices. Such practices consist of activities that are themselves rooted in each creature's *practical stance* of reciprocal authority and responsibility to all others in its community.[1] Such stances are overtly enacted and then, over time, socially instituted in a *wholly non-miraculous* fashion, by *natural* creatures.

Moreover, the resulting discursive/conceptual activities are open to assessment in respect of truth, accuracy, correctness, and overall fidelity to "the facts on the ground" (as the assessor sees them). Indeed, the project of normative pragmatics is so obviously *not super*natural that it's not clear why the self-avowed "naturalist" should be at all worried. Even less clear is why *anyone* should get to dictate the terms of legitimate inquiry *a priori*. Why, after all, should our metasemantic theorizing not make any use of the perfectly familiar and logically well-behaved deontic modal notions? Indeed, why even a *lesser* use of them than their alethic counterparts? What's so special about *alethic* modality, anyway? Nothing much, so far as I can see.

Let me dwell on this point, for it seems to me that the knee-jerk resistance to normative theorizing is deeply ingrained in the naturalist's mind. (I should know!) Pressing back against what now strikes

[1] There's been confusion on this point, caused largely, I think, by Brandom's uncharacteristically ill-chosen terminology. He gives the label "practical attitudes" to what I've here (re-)described as "embodied practical stances" on the part of a creature toward its community members. The use of the term 'attitude' has predictably conjured in the minds of some critics the notion of a propositional attitude—something that already bears a distinctively conceptual/intentional content. Plainly, this would render Brandom's account viciously circular, as he is aiming to explicate the notion of propositional attitude content in terms of (what he calls) practical attitudes. If the latter already have intentional contents, then there's obviously no difficulty in spelling out other semantic/intentional notions downstream. Equally obviously, there would be no theoretical interest in doing so.

me as an irrational prejudice, I exhort philosophers to actively discourage it, whatever the fate of Brandom's philosophical project—or, for that matter, mine—turns out to be. Given our daily immersion in social norms and institutions, it's frankly puzzling that so many theorists have allergic reactions to a deontic treatment of language. Norms are *not* puzzling. They are all around us, every moment of our lives. They permeate every social interaction we have and they are the subject of most of our thoughts, *all* of our plans, and our very conceptions of our own identities as free, responsible agents.

Moreover, with respect to linguistic norms in particular, there are (so far) no *obvious* examples in the natural world of linguistic abilities arising in creatures outside of a relatively elaborate social context. Indeed, even intelligent artifacts wouldn't count, if we ever made one, for they'd be related to the human community of happy roboticists in an obviously relevant way. So it's not at all clear—not to me, at least—why this aspect of naturalism should constrain our inquiries into language and mentality. Obviously, naturalism has many *other* appealing features, but this doesn't seem to be one of them.

The deafening silence from classical naturalists on this point has led some, e.g., Price (2010, 2013), to endorse Brandom's normative inferentialist project and to embed it into a larger philosophical framework that eschews notions of correspondence, reference, "the representation relation", etc. altogether. (No "mirrors", he enjoins, using Rorty's metaphor.) Price applies Brandom's expressivist account of logical vocabulary to *all* of human language. The resulting "global expressivism" is a key commitment of the novel brand of naturalism that he recommends to our attention—one that I find deeply compelling.

What I've been describing as the "traditional" or "classical" naturalist view—i.e., the received view among *soi-disant* naturalists in the literature from Fodor onwards—maintains that we should draw on the tools, models, and concepts of natural science in characterizing atomistic word-object or sentence-fact relations—paradigmatically, reference and correspondence. On this picture, "the world" is seen through the lens of natural science (a category itself subject to some dispute). This is a metaphysical framework that has plenty of room for protons, genes, and brains, but stubbornly refuses to accommodate responsibilities, entitlements, and the like—including, remarkably, *persons* (at least not in the fullest sense of that word; cf. Sellars 1963). The "objects" to which language relates us are thus limited by naturalist maxims to only the "natural" ones—whatever *those* are. For this reason, Price calls this view "object naturalism".

The alternative that Price puts on offer is *subject* naturalism. This view retains a healthy and well-deserved respect for the deliverances of natural science, but refuses to go along with the *philosophical* fiction of "naturalizable" reference and correspondence relations. Rather, our naturalistic urges should be directed, Price argues, toward concept- and language-using *subjects*—i.e., the creatures who acquire, produce,

and consume languages, as just one tool in a larger biological-*cum*-social enterprise of maintaining homeostasis in the species. Paradigmatically, such creatures are human persons, but any other naturally social creature can in principle be studied in this fashion.

What's striking, to my mind, is how similar all of this sounds to the methodological aims of theorists like Chomsky and Pietroski. Although both call themselves naturalists, each has made determined efforts to debunk the idea that word-world relations are relevant to an empirical study of the human language faculty. Nor does either theorist harbor the ambition—characteristic of the "classical" naturalists mentioned earlier—to *reduce* intentionality, either by analysis or metaphysically, to some alethic modal base. Here's Pietroski on the issue:

> One can raise further questions about the sense(s) in which *Begriffsplans* are intentional, and which philosophical projects would be forfeited by appealing to unreduced capacities to generate and execute the instructions posited here; cp. Dummett. But my task is not to reduce linguistic meaning to some nonintentional basis. It's hard enough to say what types of concepts can be fetched via lexical items, and which modes of composition can be invoked by complex Slang expressions.

This is another point of convergence with Brandom's pragmatism, which likewise renounces the reductive aims of the classical naturalist project.

Of course, the mere fact that Pietroski declines to take up the issue in *CM* doesn't mean he has no dog in the fight elsewhere. (I don't, myself, know.) By the same token, although it's true that Brandom doesn't aim to reduce intentional notions to some construct of natural science, it doesn't follow, and isn't true, that he has *no* reductive ambitions at all. To the contrary, his normative inferentialism is designed precisely to reduce intentionality to *something* nonintentional, which, in his case, happens to be the normative. This is why normative pragmatics serves for him as a *meta*semantics, in the fullest sense of the word. The 'meta' indicates not only that what's on offer is a "theory of meaning"—rather than a first-order "meaning theory", to use Dummett's distinction. More importantly, it connotes that the semantics is herein *subordinated* to (i.e., must "answer to", in an explanatory sense) the social norms that are the centerpiece of the pragmatics.

3.3 *Referential purport*

In keeping with his commitment to the methodological tenets of individualism and internalism, Pietroski applies many of the points scouted above to *conceptual thought*.

> [S]ome readers may be unfamiliar or uncomfortable with talk of using representations to think about things in certain ways. So some clarification, regarding aboutness and ways, is in order. … The relevant notion of thinking about is *intentional*. We can think about unicorns, even though there are none. One can posit unicorns, wonder what they like to eat, diagnose various observations as symptoms of unicorns, etc. Similarly, one can hypothesize that some planet passes between Mercury and the sun, call this alleged

> planet 'Vulcan', and then think about Vulcan—to estimate its mass, or wonder whether it is habitable—much as one might posit a planet Neptune that lies beyond Uranus. An episode of thinking about something can be such that for each thing, the episode isn't one of thinking about that thing. … [I]n the course of proving that there is no greatest prime, it seems an ideal thinker can at least briefly entertain a thought with a singular component that purports to indicate the greatest prime. … Paradoxes like Russell's remind us that even idealized symbols can fail to make the intended contact with reality, at least if one is not very careful about the stipulations used to specify interpretations for the symbols.

I want to highlight a key point here: Pietroski is presupposing about concepts *not* that they *succeed* in referring—though he allows that some of them do—but that even *empty* concepts have intentional contents. These latter plainly *cannot* be accounted for by positing straightforward metaphysical relations between words, on the one hand, and bits of the world, on the other.

This emphasis on intentionality in the sense of referential purport is crucial to Brandom's project as well. Rather than setting out to explain *successful* reference/denotation, as the paradigms of perceptual and demonstrative reference have led many theorists to do (2.3), Brandom sees it as necessary to first explain how a creature can so much as *purport* to refer to one thing rather than another, and only later to furnish an account of what counts as success in this regard. Brandom is not alone in adopting this strategy. In the Gricean tradition, the homologous project is cast in terms of the *intentional design* of communicative acts—in particular, a speaker's intentions to refer, denote, predicate, or to speak truly of something. But whether one uses the idioms of purport, design, or intention, the key point is that the phenomenon under discussion does not involve a unique, naturalizable, or semantically-relevant mind-world relation.

That leaves wide open issues about the interface between language and reality, let alone larger questions of metaphysics and epistemology. While Pietroski stays largely neutral on such topics in *CM*, Brandom's account makes quite definite commitments in these arenas. Nevertheless, there is one place where the two quite clearly converge, and that is with respect to their treatments of *de dicto* and *de re* attitude ascriptions. Let's take a look at that.

3.4 *De dicto and de re constructions*

In granting reasonable concessions to philosophers who stress the importance of mind-world relations for our theories of intentionality, meaning, concepts, and the like, Pietroski makes the following remarks:

> I grant that 'think about' can also be used to talk about a relation that thinkers can bear to entities/stuff to which concepts can apply; see, e.g., Burge. In this "*de re*" sense, one can think about bosons and dark matter only if the world includes bosons and dark matter, about which one can

> think. Any episode of thinking *de re* about Hesperus is an episode of thinking *de re* about Phosphorus. This extensional/externalistic notion has utility, in ordinary conversations and perhaps in cognitive science, when the task is to describe representations of a shared environment for an audience who may represent that environment differently. But however this notion is related to words like 'think' and 'about', many animals have concepts that let them think about things in ways that are individuated intentionally. We animals may also have representations that are more heavily anchored in reality; see, e.g., Pylyshyn. But a lot of thought is intentional, however we describe it. (80)

It's important to see that there are two distinct strands of thought here. One is about how *some* representations—again, paradigmatically those involved in perceptual or demonstrative reference—are "more heavily anchored in the world" than others. This might *seem* to put Pietroski's view at odds with Brandom's inferentialism, given how small an explanatory role the latter gives to such "anchoring" relations. But this is worry is spurious.

Brandom's account of perceptual commitments and default entitlements (*MIE*, ch. 4), as well as his (largely independent) account of demonstrative reference and "object-involving" thoughts, are fully compatible with—indeed, positively *require*—the existence of reliable nomic or causal relations between perceptible objects in the world and the perceptual mechanisms of a creature. There are, to be sure, heated debates about how exactly all of that works—e.g., whether the percepts should be seen as having the function of bare demonstratives (Fodor and Pylyshyn 2015), noun phrases (Burge 2010), or inferentially integrated singular thoughts (Brandom 1994). But disagreements on these points are far downstream, theoretically, from the broadly methodological commitments that I want to highlight here. It's these that are the subject of the second strand of thought that I think we should distinguish here.

When Pietroski speaks of a "'*de re*' sense" in which one can think about or ascribe thoughts, he is at once talking about a certain kind of *thought*—the "object-involving" kind discussed above—but also, separately, about a certain kind of thought *ascription*. The latter, he says, "has utility, in ordinary conversations and perhaps in cognitive science, when the task is to describe representations of a shared environment for an audience who may represent that environment differently" (*op. cit.*) Note that this circumscribes the function of *de re* ascriptions to what I think of as "the navigation of perspectival divides". More prosaically, *de re* constructions allow language users to describe the environment in which a creature is deploying its concepts—as viewed *by the describer* (and often her audience)—while *de dicto* constructions function to ascribe *the concepts so deployed*. Here, then, we see another major point of convergence between Pietroski and Brandom. The latter provides an inferentialist analysis of *de dicto* and *de re* ascriptions, according which they perform precisely the function that Pietroski's remarks indicate.

According to Brandom, *de dicto* ascriptions make explicit the commitments of the creature being described—not those of the ascriber, who may be either ignorant on the matter or hold commitments directly contrary to those ascribed. A speaker might say, "Dan thinks the greatest philosopher of language was Quine," without having any commitments one way or the other about whether philosophers exist, or about whether Quine was one of the greats. Indeed, the speaker might think that philosophers are bits of tofu and that Quine is a particularly flavorful brand. None of that would matter with regard to the speaker's entitlement to a *de dicto* claim about what *Dan* said.

By contrast, had the speaker employed a construction that functions in the *de re* sense—e.g., the awkward 'of'-constructions that we've inherited from Quine—then their own commitments would have come into play, with questions arising (at least in potentia) about their *entitlements* to those commitments. For instance, had the speaker's ascription been (1), then their own commitments regarding the existence of philosophers, the list of the greats, and the possibility of philosophical tofu would have become immediately relevant.

(1) Regarding the greatest philosopher of language, Dan thinks that *he* was a piece of tofu!

Turning to subsentential cases, Brandom points out that such shifts in perspectival commitments can be indicated by operators such as "classified as", "described as", "conceptualized as", and (importantly) "referred to as". For instance, in saying "Jamal classified some food *as rabbit*," a speaker, Juanita, purports to indicate some food—the *de re* component of the ascription—and then says what concept Jamal applied to it (viz., RABBIT). The word 'as' marks the onset of the *de dicto* component, ensuring that Juanita does not commit *herself* to the correctness of Jamal's classification. (Perhaps she knows that that the stuff on the plate is tofu.)

Some theorists, having noted that RABBIT is the *only* concept that Juanita ascribes to Jamal, go on conjecture that Jamal can deploy this subsentential concept, *all by itself*, in classifying something as rabbit. Indeed, bewitched by surface grammar, some fail to notice the plain distinction between *what Juanita is doing*—i.e., describing (noncommittally) one *aspect* of Jamal's perspectival commitments—and *what Jamal is thinking*. A mistaken conflation of these two phenomena is what gives rise, I suspect, to the widespread illusion that we can take as it *as a datum* that each of us has ability to think of, classify, or refer to something, by deploying *just one subsentential concept* (or linguistic expression). I'll argue later, following Brandom, that this idea is doubtful; *prima facie*, one can neither classify nor think about something without tokening *complete thoughts*. For instance, the case described, Jamal's classificatory act requires tokening the *complete* thought, CCC IS-RABBIT, where CCC is whatever concept he uses in thinking about the food on the plate.

3.5 *Interpersonal similarity of meaning and content*

Many theorists hold, for reasons that are ultimately unpersuasive, that communication is a matter of passing a message, idea, meaning, or thought from speaker to hearer. As with throwing a frisbee, a successful case in one in which the item sent by the speaker is the very same as the one received by the hearer—if only in the ideal. But, after the mid-century work of Quine, Kuhn, Sellars, and many others who developed broadly holist ideas (e.g., Churchland 1979), it's hard to see this picture as anything but optional. Brandom's account of communication is one of the many that rejects it outright, modeling communication instead on activities like tango dancing, where partners has to give and take "in complementary fashion"; book-keeping, where each participant "keeps separate books" regarding her own commitments and those of other participants; and baseball games, where a common "scoreboard" shows what complex normative statuses each participant bears to each of the others.

Metaphors aside, Brandom's inferentialism carries an explicit commitment to *holism* about meaning and content. Many follow Fodor and Lepore (1992) in seeing this as the root of several major problems for his view. But the objections that Fodor and Lepore press are virtually all rooted in implausibly strong assumptions about the necessity of meaning/content *identity*—rather than mere similarity—for various philosophical inquiries. These include the projects of adequately characterizing successful communication, interpersonal disagreement, and rational belief revision. It seems to me that Brandom's accounts of these things are perfectly fine as they stand, but fans of holism seem to have gone scarce in recent years, and Fodorian views about meaning/content identity have arguably become the *received* views in the field.

In yet another clear case of both his iconoclasm and his significant convergence with Brandom, Pietroski likewise rejects the identity-of-meanings picture, though on grounds that are independent of any holist commitment. Targeting first the extensional account of meaning-identity as co-extensiveness, Pietroski points out that

> speakers can connect the pronunciation of 'pen' to the same meaning without determining a set of pens that they are talking about. If each speaker uses a polysemous word that lets her access a concept of writing implements, as opposed to animal enclosures, they resolve the *homophony* the same way and thereby avoid blatant miscommunication. In this sense, they connect a common pronunciation with a common meaning. But it doesn't follow that any speaker used a word that has an extension... (18)

Later, Pietroski counsels us—wisely, in my view—to give up the whole idea that "successful communication requires speakers [to] use the same meanings/concepts" (33), regardless of the theoretical framework in which this idea is couched (extensional or otherwise). He views it as a mere *idealization* that "members of a community have acquired the *same* language—or that they use the same words, or words that have the same meanings." Despite Brandom's focus on social norms, shared

commitments, and the like, what Pietroski says here is entirely in line with his view. This will take some spelling out, which I undertake in 4. For now, let me emphasize that, if correct, then Pietroski's take on this issue would severely undermine the (already fairly flimsy) arguments against meaning holism.

3.6 *The Pragmatics of Assessment*

We've seen that Brandom places great stress on the notion of assertion, and that he sees this as something that we should characterize in normative terms. Given Pietroski's naturalist commitments, one might think that he disagrees. But this overlooks the key point that I will go on to make in the remainder of this discussion—namely that the theoretical aims of Pietroski's semantic theory are so starkly different from those that animate Brandom's inquiry that their common use of folk terms like 'meaning' should not bewitch us into thinking that they're talking about the same phenomenon.

Pietroski's target throughout *CM* is not communication, but the underlying psychological mechanisms that make it possible. When he does provide hints of his broader views about communication, what he says is entirely of a piece with Brandom's normative pragmatism.

> My own view is that truth and falsity are properties of certain evaluative actions—e.g., episodes of assertion or endorsement—and the corresponding propositional contents that are often described with the polysemous word 'thought', as in 'the thought that snow is white' or 'the thought constructed by executing that instruction'; cp. Strawson. (63, fn. 23)

As with all theory-laden terms, including (especially in this context) 'thought', 'meaning', and 'concept', we should always remember that the aims and presuppositions of the inquiry are far more important to keep in view than the pronunciation of the jargon. Pietroski makes this point in the following passage.

> Let's not argue about nomenclature. One can use 'concept' more permissively than I do, perhaps to include images or other representations that are not composable constituents of thoughts. One can also use 'concept' less permissively, perhaps to exclude representations that fail to meet certain normative constraints. Or one might reserve the term for certain contents, or ways of thinking about things, as opposed to symbols that have or represent contents. But I want to talk about a kind of physically instantiated composition that is important for cognitive science, along with a correspondingly demanding though non-normative notion of constituent, without denying that contents and nonconceptual representations are also important. (77)

We will go on in 4 to compare Brandom's and Pietroski's notions of "constituent," where it will become important that Pietroski's aims is to provide a *non*-normative account of constituency. For now, consider how this plays out with respect to Pietroski's distinction between Olde and New Mentalese (2). Recall that Pietroski's account has children starting off with Olde Mentalese concepts that fail to meet the conditions for assembly by FL *Begriffsplans*, and later *inventing* new concepts that are

specially tailored for the job. For instance, Pietroski writes, "Ignoring tense for simplicity, we can form concepts equivalent to: GIVE(VENUS, __ , BESSIE); GIVE(VENUS, BESSIE, __); GIVE(__, BESSIE, VENUS); GIVE(__, BESSIE, __); etc. But children may start with composable concepts that are *less* cognitively productive" (101).

In the remainder of the discussion, I'll lean heavily on the following two key points. First, much of what Pietroski says about the phylogenetically older "Fregean" modes of thought that permit internal substitution is of a piece with the lessons that Brandom likewise draws from Frege. The second key point, which alone serves to resolve many of the apparent conflicts between the two views, is that Pietroski's account is best be viewed as a (partial) theory of the *subpersonal* mechanisms that underlie our norm-tracking abilities, which in turn contribute to making possible the social practices of communication. For instance, Pietroski's account of how a child comes to acquire the concepts that FL can assemble should, if correct, be regarded as laying down empirical constrains on the variety of norm-tracking social practices that a human child will master at various ages.

If the foregoing claims are correct, then the two theoretical enterprises not only share several key commitments, but they are actually *complementary*, each providing an important piece of the overall puzzle about how best to view language(s). While Brandom's inquiry is neither descriptive nor psychological, his account clearly presupposes that there must be *some* descriptively correct account of how any creature's sub-creatural cognitive architecture emerges, whether in early development or in its species' socio-evolutionary ("memetic") history. His claim is simply that such an account can't be the *end* of the story. Rather, he argues, it's the necessary groundwork for a much larger picture of the role of language in social practice.

Likewise, although Pietroski's focus is the psychology of the individual speaker-hearer—assumed to be universal within the species—we've seen that there's room for communal norms in his overall picture. Indeed, passages from *CM* contain explicit remarks about the norms that govern inquiry in the mature sciences—e.g., norms that serve to stabilize referential purport among expert chemists. Such norms, however, would reside largely at the level of *pragmatics*, which is decidedly not Pietroski's focus in *CM*, nor a major aspect of Chomsky's own work.

Perhaps, then, we can adapt David Marr's familiar distinction between "levels of analysis"—if only by crude analogy—by viewing Brandom as articulating a "high-level description" of the conditions that any language user must satisfy in order to count as such. If this suggestion is right, then Pietroski is best seen as telling us about the details of how this high-level description happens to be "implemented" or "realized" in the human case. In the remainder of this essay, I will argue that there is no inherent conflict between accounts pitched at distinct levels of inquiry. There could only be a substantive dispute if the they

shared a common domain. But, as Pietroski's remarks in the following passage make clear, that's just not so in the present case.

> Let me digress, briefly, to note a different response to concepts without *Bedeutungen*. One might adopt a hyper-externalistic understanding of 'ideal', so that no ideal thinker has the concept VULCAN, and no ideal analog of this concept lacks a *Bedeutung*; cp. Evans, McDowell. But then an ideal thinker must not only have her mental house in good order, avoid paradox-inducing claims, and be largely *en rapport* with an environs free of Cartesian demons. In this hyper-externalistic sense, ideal thinkers are immune from certain kinds of errors, perhaps to the point of being unable to think about the same thing in two ways without knowing it; cp. Kripke. One can characterize corresponding notions of concept*, language*, and their cognates. These notions, normative in more than one way, may be of interest for certain purposes. Perhaps inquirers aspire to have languages* whose expressions* have meanings* that are instructions to build concepts* that can be constituents of thoughts*. But my aim, more mundane, is to describe the thoughts/concepts/meanings that ordinary humans enjoy. And I see no reason to believe that the best—or even a good—way to study concepts or meanings is by viewing them as imperfect analogs of their hyper-externalistic counterparts, even if we can and should try to acquire concepts*. Notions like truth and denotation may be required to interpret philosophical talk of thoughts*/concepts*/meanings*. One can also hypothesize that ordinary words combine to form sentences that express thoughts composed of concepts that denote or "are true of " things in the environment. But this hypothesis does not become inevitable as soon as we talk about thoughts/concepts/meanings. On the contrary, natural concepts seem to be mental symbols that can be used to think about things, even when the concepts apply to nothing. And if words like 'give' are used to access naturally introduced concepts like GIVE(_), we must be prepared to discover that words are used to access concepts that fall short of certain philosophical ideals.

4. *Challenges for Ecumenicism*

The ground-clearing maneuvers of the previous section put us in a position to explore residual differences between Pietroski and Brandom that threaten to be more substantive. As advertised, I hope to show that the initial appearances are misleading even in some of *these* cases, contrary to received opinion amongst philosophers of language. Still, I'll ultimately admit defeat when we arrive at the very last topic—the banes of predication and singular terms.

4.1 *E-language*

One place to look for sources of substantive disagreement is in the vicinity of Chomsky's infamous distinction between E-language and I-language (Chomsky 1986). Chomsky initially drew this distinction with the explicit intention of formulating a key difference between his approach to language and that of Quine, Lewis, and others—what was then arguably the dominant view. Given that Brandom was a student of Lewis's and frequently pays homage to him in published work, the

question arises whether his view likewise falls prey to the compelling objections that Chomsky articulated decades ago. We can address this question by focusing on Pietroski's more recent formulation of Chomsky's insights.

Let's begin with the distinction itself. What exactly is an E-language supposed to be?

> Chomsky characterizes I-languages as procedures that connect interpretations with signals. Languages in any other sense are said to be E-languages. Chomsky … isolates a procedural notion and uses 'E-language' as a cover term for anything that can be called a language but isn't an I-language. Thus, E-languages may include certain clusters of behavioral dispositions, heuristics for construing body language, etc. But Chomsky's 'I/E' contrast does connote Church's intensional/extensional contrast, and sets of interpretation-signal pairs are paradigmatic E-languages. Though to repeat, such sets are often defined procedurally. (52)

This negative conception of E-languages ("anything that can be called a language but isn't an I-language") casts such a wide net that it *can't but* apply to Brandom. Certainly, the latter is making no substantive psychological hypotheses about a computational system in the human brain. Nevertheless, as we'll see, much of what Pietroski says in rejecting the study of E-language poses no conflict with Brandom's view. While he argues that David Lewis "wanted to describe our distinctively human language(s)," this is *not* the direction that Brandom takes. His project is to describe language *as such*, not "distinctively *human* languages."

Although my broader claims by no means hang on a particular interpretation of Lewis, it does seem to me that Pietroski's reading of Lewis is not maximally charitable. I say this not because I think that his deep disagreements with Lewis somehow insidiously creep into his interpretation. Rather, I'll argue that there's a way of seeing Lewis as engaged in a wholly different project from the one that Pietroski foists on him. Still, whatever the case about Lewis, the project that I have in mind is the one that Brandom *in fact* undertakes—which, in my view, he carries out successfully, even if his teacher did not.

The subsections that follow distinguish three separate points of debate: (i) extensional vs. procedural conceptions of grammar, (ii) a sentence-first vs. word-first approach to semantics, and (iii) individualist vs. social conceptions of language.

4.1.1 *Extensional vs. intensional constraints*

In getting more precise about Lewis's particular brand of E-language, Pietroski highlights two points: (i) the metaphysical claim that language(s) are extensionally-specified abstract objects—specifically, sets of meaning-pronunciation pairs—and (ii) Lewis's conception of the process whereby a population comes to *use* such an object in social exchanges.

> Lewis's proposal was that languages like English are sets [functions-in-extension, not procedures] that are related to certain social phenomena via conventions of "truthfulness and trust." He speaks in terms of populations "selecting" certain languages. The suggestion is that using a particular language is like driving on the right: an arbitrarily chosen way of coordinating certain actions. (55)

One of the main things that Pietroski finds problematic here is that treating natural languages as functions-in-extension requires relinquishing the explanatory ambitions of generative linguistics—specifically, the aim of describing the cognitive structure that underlies, or even *constitutes*, human linguistic competence. Lacking an account of the cognitive architecture of FL, or the representational format in which it conducts its business, we have no empirically credible story about how adults accomplish real-time language processing, not to mention how children acquire the creative capacity to produce and consume an indefinite range of novel expressions.

> Lewis says instead that "a grammar, like a language, is a set-theoretic entity which can be described in complete abstraction from human affairs." Chomsky offers a way of locating languages and grammars in nature. Lewis stipulates that each grammar determines the language (i.e., the set of sentences) that it generates, but that languages do not determine grammars. So even if there is a Lewisian grammar for a certain set of Slang sentences, this does not explain how the relevant sentence meanings are related to lexical meanings, or how speakers of the Slang know that the strings in question fail to have certain meanings. At best, a Lewisian grammar indicates how a certain kind of mind might abstract lexical meanings from sentence meanings, given hypotheses about the relevant constituency structures and composition principles; see chapter three. But this doesn't yet tell us anything about how humans connect lexical meanings with pronunciations, or how we can/cannot combine lexical meanings. (59)

This all strikes me as correct. What Lewis *should* have been aiming his inquiry at language *as such*, regardless of which creature is using, acquiring, or "selecting" it. Setting aside Lewis, Brandom's target explanandum is "what it is to do the trick," not "how the trick is done by *us*"—or, for that matter, by a dolphin, an AI robot, or a Martian. (I'll suggest a friendly amendment to this point in 4.2). Chomsky, Pietroski, and other generativists, on the other hand, want to know about humans *specifically*. These are different projects, to be sure, but they're not thereby rivals. Though they may well constrain one another, the relations between them can better be seen in terms of the "levels of analysis" picture that I proposed earlier.

In addressing Lewis's account of how populations "select" a language, Pietroski assumes that the target phenomena are sufficiently similar between his inquiry and Lewis's that the two are not only commensurable with one another, but are actually in direct conflict on various key points.

> Lewis asserts that £ is the language used in a population P "by virtue of the conventions of language prevailing in P;" where conventions, in his sense,

> are special cases of mutually recognized regularities of action/belief in a population of individuals who can act rationally. He says that a convention of " truth and trustfulness," sustained by " an interest in communication," is what makes a particular set of sentences the language of a given population. However, Lewis offers no evidence that this proposal is correct. And there is an obvious alternative: if a population P is small enough to make it plausible to speak of the language used in P, then the members of P will have acquired generative procedures that are very similar. So why think there are "facts about P which objectively select" a shared E-language, and not that members of P have similar I-languages? (60)

The answer that I suspect Lewis would give to Pietroski's last question in this passage is that there is no obvious way of individuating I-languages independently of the "forms of life"—i.e., communal practices—in which any linguistic creature is caught up. That's because a discursive creature's concepts and/or lexical items are constitutively related to the norms that structure those practices. Thus, if meanings really *are* instructions to fetch and assemble concepts, then we won't be able to individuate *meanings* without appealing to such norms *either*. It's not clear to me how Pietroski would (or could) respond to this point.

Pietroski's critique of Lewis—particularly, his troublesome notion of "selection"—is carried forward in the following passage. What's new here is that we have a direct quotation from Lewis making precisely the claim that I've been urging on his behalf—namely, that the generative grammarian's target *explananda* are simply *not the ones that* he *seeks to address*.

> According to Lewis, populations select languages by virtue of using sentences in rational ways. So on his view, if language use is not a "rational activity" for young children, they are not "party to conventions of language" or "normal members of a language-using population." [Lewis writes:] "Perhaps language is first acquired and afterward becomes conventional. . . . I am not concerned with the way in which language is acquired, only with the condition of a normal member of a language-using population when he is done acquiring language." I don't know what it is to acquire language in Lewis's sense, or how he would describe whatever creolizers acquire. But even if one wants to focus on "normal" adults, ignoring acquisition may not be a viable option for those who want to find out what Slangs and meanings are. Inquirers don't stipulate how phenomena are related; they investigate.

In the end, I agree with Pietroski that Lewis's talk of populations "selecting" languages is hard to take seriously, especially in light of the Chomskyan alternative. Casting things in Lewis's way plainly does leave the crucial process of "selection" unexplained—at least at the level of psychology, which is arguably where all the exciting action is at. Nor can I bring myself to credit the idea that the central aim of formal linguistics should forever be what Lewis said it should be—i.e., the extensional characterization of some abstract set (a moving target, as Pietroski points out, given the constant introduction of novel lexical items). But one can share Pietroski's doubt that there's any meaningful sense in which people "select languages", as well as his broader anti-

extensionalism, but *nevertheless maintain* that there's a level of theoretical abstraction at which language *is* best viewed as a distinctively social phenomenon.

The conventions that govern the phenomenon in question might be "truthfulness and trust", as Lewis thought, or they might be more explicitly normative, as in Brandom's conception of socially-instituted assertional and inferential commitments and entitlements. Likewise, the languages that such conventions institute might be static abstracta, as Lewis seems to have believed, or they might be flexible, dynamic, and highly context-dependent social relations, as on Brandom's account. The latter, moreover, has no theoretical use for the notion of "stable populations" that go around "selecting" various abstract objects—each of which somehow manages to answer to the label 'Spanish' (or whatever), despite their extensional differences.

There is, therefore, no call to reify such strange entities on the normative inferentialist picture. Neither linguistic phyla (e.g., Romance or Germanic) nor local dialects (2020 Boston English) need be real—not in *Lewis's* sense, anyway—in order for Brandom's project to get off the ground. Indeed, as previously mentioned, the "communities" that Brandom appeals to in his account of norm-governed social practices can be as small as *two creatures*. (Actually, a footnote in *MIE* reveals that Brandom might even be willing to countenance something like an "I-thou relation" coming to be instituted by temporally asymmetric recognition of authority and responsibility relations between *distinct time-slices of one and the same creature*.) Thus, in principle, every new dyadic social interaction can serve to institute novel *local* norms, alongside any that were previously shared. Whatever else Brandom might be accused of doing, then, attempting to individuate and reify stable (let alone timeless) public languages is no part of his brief.

4.1.2 *The primacy of sentences*

Having argued against Lewis's *metaphysical* assumptions, Pietroski goes on to take issue with his *methodological* claim that semantic inquiry should begin by assigning meanings to sentences, rather than to subsentential expressions.

> Lewis goes on to say "if σ is in the domain of a language £, let us call σ a sentence of £." But he wasn't using 'sentence' as a technical term for any string to which a language assigns a meaning. Rather, Lewis initially restricted his notion of a meaningful expression to sentences. He later introduces talk of word meanings via talk of grammars. (55)

As we saw in our earlier discussion of Pietroski's views on sentences, he views this sort of approach as being out of step with contemporary theorizing in generative linguistics.

> Linguists have since replaced "S" with many phrase-like projections of functional items that include tense and agreement morphemes, along with various complementizers. This raises questions about what sentences are, and

> whether any grammatical notion corresponds to the notion of a truth-evaluable thought. But theories of grammatical structure—and to that extent, theories of the expressions that Slangs generate—have been improved by *not* positing a special category of sentence. So while such a category often plays a special role in the stipulations regarding invented formal languages, grammatical structure may be independent of any notion of sentence. (61)

Here again, we must distinguish Lewis from Brandom. For the latter, the primacy of sentences for is not a mere stipulation, nor an irrational fetish for a specific syntactic type. Rather, his principal—and I think quite principled—grounds for isolating sentences at the outset of a normative pragmatic inquiry is that this is the only type of expression with which a can make an *assertion*—i.e., an explicit move in an norm-governed *inferential* practice. Brandom's concern with the normative structure of this "game of giving and asking for reasons" is what motivates his focus on the role that assertions play, both in reasoning—as premises and conclusions—and also in communication. In the latter context, they serve the function of allowing speakers to undertake normative statuses—paradigmatically, commitments and entitlements.

Pietroski's methodological counsel is to postpone discussion of communication to a later day. As he remarks, specifying the structure of FL is work enough for one lifetime—surely more. But it doesn't follow from this that a theoretical inquiry with *different* aims must be, in some sense, second-rate, let alone illegitimate. Nor is it clear that the two inquiries are even in competition with one another over how to best to describe a common subject matter. As noted above, their common talk of "meanings", "languages", and the like might tempt one into thinking that the topic under discussion is the same for both theorists. But that would be a mistake. The two theoretical frameworks in which these terms are couched—and relative to which they have their theoretical import—are so dramatically different in respect of their explanatory aims that viewing them as even *purporting* to refer to the same phenomena is a bit of a stretch.

Unless I'm mistaken, Pietroski himself seems to have fallen into this trap, despite his earlier counsel to avoid such temptations. One clear place where he does so is in the following passage, which is notable for including—now for a second time—Lewis's direct protests against being saddled with the views that Pietroski nevertheless goes on to attribute to him.

> By contrast, Lewis held that sentences are prior to grammars. While granting that we should not "discard" notions like phrasal meaning (relative to a population P), or the "fine structure of meaning in P of a sentence," he says that these notions "depend on our methods of evaluating grammars." … For Lewis, a grammar Γ is used by P if and only if Γ is a best grammar for a language £ that is used by P in virtue of a convention in P of "truthfulness and trust" in £. One might have thought that a "best" grammar for the alleged set £ would be one that best depicts the procedures acquired by members of P. But according to Lewis, "it makes sense to say that languages might be used by populations even if there were no internally represented gram-

> mars." He then makes an even more remarkable claim. "I can tentatively agree that £ is used by P if and only if everyone in P possesses an internal representation of a grammar for £, if that is offered as a scientific hypothesis. But I cannot accept it as any sort of analysis of "£ is used by P", since the analysandum clearly could be true although the analysans was false." Note the shift from Lewis's opening question—what is a language?—to a search for an analysis of what it is for a language to be used by a population. … This shift is interwoven with the insistence that languages are sets of sentences, and a willingness to accept the consequences for how grammars are related to non-sentential expressions.

What Pietroski describes here as a "shift" in Lewis's explanatory aims strikes me as a correct description of what Lewis was up to all along; or, at any rate, *should've* been.

Setting aside Lewis exegesis, we can turn again to Brandom, whose methodology is a good deal more perspicuous. As already noted, Brandom rejects the "set of sentences" conception of languages, tying social norms to practices that are fluid, socially distributed, and highly context-dependent. Nevertheless, he can accept Lewis's views (quoted in the passage immediately above) about grammars and about the status of subsentential expressions. Again, that's because he motivates the primacy of sentences (or "full thoughts") by reference to their roles in assertions, and hence in the norms governing the social practices in which assertion is a basic move. Such a practice need not be as articulated as ours, in respect of either content *or* syntax. And while Pietroski is prosecuting an empirical inquiry into *human psychology*, Brandom's is assaying the pragmatics of *assertion*, *inference*, and *assessment*.

Here is one final bit of textual evidence for my suggestion that Chomsky and Pietroski, on the one hand, and Lewis and Brandom, on the other, are simply talking at cross purposes. Pietroski writes:

> I think [Lewis's] ordering of priorities is misguided. Slangs are child-acquirable generative procedures that connect meanings with pronunciations in ways that allow for constrained homophony. So whatever meanings are, there are natural procedures that connect them with pronunciations in specific ways. Instead of adopting this promising starting point for an empirically informed discussion of languages and meanings, Lewis offered a series of stipulations. Many others followed suit. But that doesn't make it plausible that Slangs are sets. And if Slangs are I-languages in Chomsky' s sense, then we shouldn't ignore this fact when asking what meanings are. … [Lewis] adds that "the point is not to refrain from ever saying anything that depends on the evaluation of grammars. The point is to do so only when we must, and that is why I have concentrated on languages rather than grammars". But in reply to a worry that he is needlessly hypostatizing meanings, he says "There is no point in being a part-time nominalist. I am persuaded on independent grounds that I ought to believe in possible worlds and possible beings therein, and that I ought to believe in sets of things I believe in." So why be a part-time grammarist, given Chomsky's reasons for thinking that children acquire generative procedures? Quine's worries about the "indeterminacy" of meaning are not far away. But while Lewis speaks of evaluating grammars, he does not engage with Chomsky's notion

> of an evaluation metric, or the correlative notion of "explanatory adequacy".

I suspect that the reason Lewis would give for being what Pietroski disparagingly calls a "part-time grammarist" is that he is—as he says repeatedly—*not* concerned with specifically *human* languages, but with language *as a general phenomenon*. This, in any event, is the line that Brandom takes. And while Pietroski's arguments in favor of his semantic proposal are compelling in the context of *his* particular brand of inquiry, it's difficult to make sense of his idea that a different "ordering of priorities" can be "misguided". At worst, an ordering of priorities—i.e., the adoption of some concrete set of methodological maxims, descriptive aims, and explanatory ambitions—can fail to be illuminating about the domain that it carves out for itself. But it's hard to see how this charge can be credibly leveled against Brandom's inquiry, the results of which many philosophers—the present author included—find deeply revelatory of communicative linguistic practices.

4.1.3 *Public languages, dialects, and I-languages*

Another context in which the notion of E-language has been invoked is to account for how folk terms like 'Italian' and 'Swahili' manage to pick out something in the world. Pietroski takes a dim view of this theoretical aim.

> we can describe many I-languages as English dialects (or idiolects), without English being any particular language. Prima facie, there are many ways to be a speaker of English: American, British, and Canadian ways; young child ways, adult scientist ways; etc. Being a speaker of English seems to be a *multiply realizable* property whose instances are similar in ways that matter for certain practical purposes. We can use 'English' to group certain I-languages together, perhaps in terms of paradigmatic examples, an intransitive notion of mutual intelligibility—think of Brooklyn and Glasgow—and historically rooted dimensions of similarity. There need not be an English language that each speaker of English has imperfectly acquired; cp. Dummett. We can use 'Norwegian' similarly, and classify both Norwegian and English I-languages as Germanic, without supposing that Germanic is a language shared by speakers of Norwegian and English. Analogies between linguistic and biological taxonomy can be preserved, whatever their worth, by thinking of specific I-languages as the analogs of the individual animals that get taxonomized—with 'Human' as the most inclusive category, and 'Indo-European' indicating something like a phylum.

Turning to Brandom, we note once more that he is not seeking to taxonomize the natural languages of the human species. His theoretical aims do not require reifying or hypostasizing the norms that govern specific discursive interactions, nor individuating the entities that allegedly answer to the names 'English' or 'Norwegian'. There seems to be no principled reason why Brandom couldn't countenance Pietroski's eminently reasonable view on these matters. Thus, when Pietroski describes a hypothetical theorist who "grant[s] that children acquire I-languages, yet maintain[s] that Slangs are E-languages that connect pronunciations

with extensions of idealized concepts," the theorist he is describing is not Brandom. If the latter were in the business of theorizing about Slangs, in particular, then he might indeed take up the hypothetical proposal that "each Slang is a social object that certain speakers acquire by internalizing a generative procedure and meeting some *further* conditions." But no part of his actual view hangs on whether "English", "British English", or "Germanic" name metaphysically real entities, let alone ones that are individuated extensionally. Nor, again, is his theory a descriptive psychological one. And it most certainly has no truck with extensions.

We can illustrate all this more clearly by considering a contrast that Pietroski draws between what he calls E-NGLISH (an E-language) and NGLISH (an I-language).

> One might describe E-NGLISH in terms of the strings that certain people *could* understand as sentences, and the meanings they *could* assign to those strings, given a certain dictionary and a suitably idealized sense of 'could'. But this presupposes that a competent speaker of NGLISH has an expression-generating procedure whose lexicon can be expanded. So we don't need to invoke E-NGLISH to say what NGLISH is.
>
> One can stipulate that if Amy, Brit, and Candice speak English, there is something that Amy, Brit, and Candice speak. But then the "thing spoken" may be a class of I-languages. One can stipulate that people share a language if and only if they can communicate linguistically. But then "shared languages" may not play any role in *explaining* linguistic communication. Sharing a language, in the stipulated sense, may be a matter of using similar I-languages in combination with other human capacities and shared assumptions about potential topics of conversation.

These, I maintain, are all claims that Brandom should be happy to accept. True, he holds that there are social norms governing the communicative exchanges of creatures who can understand and produce indefinitely novel constructions. But he should have no qualms with the claim that humans happen to do this via some subpersonal psychological apparatus. Plainly, his account must presuppose that there is at least one way of "doing the trick"—our own—though it leaves room for others.

Ultimately, as I've emphasized throughout this discussion, the *appearance* of friction is, here as elsewhere, rooted in a cluster of verbal disputes. Underlying these is the fact that Brandom's normative inferentialism, while sharing many homophonous pieces of jargon with Pietroski, is pitched at a *higher level of description*, so to speak, than a cognitive theory of NGLISH. Thus, Brandom should grant that NGLISH—the psychological apparatus that a given speaker possesses—can be specified without any recourse to E-NGLISH (if such a thing can even exists in a sense that he would countenance).

Note that, if Brandom's inferentialism is on the right track, then the "shared assumptions about potential topics of conversation" that Pietroski mentions are *partly constitutive* of the discursive norms that enter into an illuminating pragmatic account of communication. Need-

less to say, these not be the only things that enter into such an account. At the level of phonological and morpho-syntactic processes, as well as the ability to access and assemble concepts, the computational—and, ultimately, neurocognitive—explanation will surely be couched in the terms that generativists recommend. These are arguably subpersonal processes, which, again, is not the theoretical "level" at which Brandom's account is pitched.[2] A substantive disagreement can only be maintained if both inquiries have a shared target; this is, once again, not such a case.

Still, I think it must be admitted that a large part of what makes a cognitive theory of NGLISH theoretically interesting is that it enters into a larger account of how a specific creature—the human animal—is able to track, follow, articulate, challenge, reject, and revise the norms of discourse that have contingently arisen in its social *milieu*. This is not the only source of theoretical interest, of course. The biologically-realized combinatorial principles that constitute NGLISH are a marvel of neurocognitive information-processing, and should be studied by natural science as such, with the fascination that grips empirical linguists and neuroscientists alike. Nevertheless, the internal operations of this apparatus would be of no *adaptive* use to a creature if they made no contribution to the shaping of a broader class of norm-governed social activities. And, biology aside, the beauty of the mechanism is only enriched, not diminished, when we take into account the social interactions that it makes possible.

Famously, Chomsky (2016) rejects the idea that language is an adaptation specifically designed for social/communicative purposes. He entertains an alternative hypothesis to the effect that the *narrow faculty of language*—what Pietroski is here calling 'NGLISH'—was initially an aid to individual thought, making it possible for recursive structures in the mind to be composed, and thus "entertained". This hypothesis strikes many (including myself) as implausible, but even if it turns out to be true, Brandom's claims would still hold. That is, it would remain the case that the creature's newly-structured individual-level thoughts played a role in its profound re-shaping of the prevailing social norms. Had its thoughts or judgments—i.e., its normatively evaluable commitments to things being thus-and-s0—not *somehow* become entwined with broader social practices, they would not thereby have been *commitments*, whatever else they might have been. This is because, lacking a social existence, there is no normative check on what the creature had committed itself *to*. The conceptual contents of the internal states of a solitary creature would thus be underdetermined to such an extent that it may be more accurate, not to mention fruitful, to view them as *subpersonal* states—or, at any rate, as something other than judgments (*strictu dictu*). But this reeks of the kind of terminological legislation that both Pietroski and Brandom repeatedly warn against, so I'll leave the matter there.

[2] I spell out what I mean by "subpersonal" in Pereplyotchik (2017: 7.3)

Putting aside terminology, what Brandom is concerned to articulate is the general structure of pragmatic social statuses—e.g., commitment and entitlements—and how the norms that govern those statuses can be used in theorizing about the semantics of linguistic expressions in any language. The project is "general" in the sense that it designed to be applicable to *any* case of linguistic communication, not just the human case. How the norms of communication that we find in humans today might have evolved in the distant past of our particular species is, once again, no part of Brandom's explanatory target. Clearly, they emerged somehow, and the empirical story is bound to be fascinating. But all Brandom really needs to get *his* project going is the very broad claim—empirical, though perhaps only by a courtesy—that *some* natural psychological mechanisms underlie, and thus explain, in the descriptive naturalistic sense, any given creature's facility with social norms.

One might suspect that there is nevertheless a substantive *metaphysical* dispute in the vicinity. Brandom posits public linguistic norms, whereas Pietroski has no truck with public entities of any kind. But matters aren't so clear. Consider how Pietroski characterizes two hypothetical ontological views about what NGLISH really is: "We can identify NGLISH with an I-language, or perhaps a class of I-languages that differ only in small ways that are irrelevant for the purposes at hand." But what the difference could there possibly be between positing an E-language and positing "a class of I-languages that differ only in small ways that are irrelevant for the purposes at hand"? What ontological payoff, that is, can there be for a theorist who insists on I-languages *and resemblance classes thereof*, as opposed to public languages—or, at any rate, theoretically useful public linguistic entities (TUPLEs)?

Pietroski hints that the issue may have to do with other metaphysical features of E-NGLISH and NGLISH, including their *modal* properties: Whereas "a Slang seems to have its composition principles *essentially*, … E-NGLISH includes no composition principles; the set contains only string-meaning pairs, atomic and complex." The difference is that sets are individuated by their members, but "any initial list of atomic expressions can be updated." Pietroski's point, if I understand it correctly, is that the ongoing process of language change creates a moving target for semanticists like Lewis. As new lexical items emerge (or "go extinct") in the actual world, the set that such a theorist intends to pick out changes. Indeed, Pietroski quips that "[i]dentifying Slangs with sets of expressions is like identifying animals with sets of molecules, and insisting that growth be described as replacing one animal with another. Even if this metaphysics is coherent, it may not cohere with plausible biology and linguistics" (57).

It isn't clear, though, why this sort of consideration couldn't be pressed just as hard with regard to I-languages. Although Pietroski says that "a Slang seems to have its composition principles *essentially*," he plainly acknowledges that I-languages, conceived of as a psychological procedures, can *also* be updated. For instance, "communicative

failures can lead children to modify their (still modifiable) procedures for connecting meanings with pronunciations, subject to constraints" (58). Indeed, it seems to be an empirical hypothesis whether all of the elements of Pietroski's *own* cognitive/semantic proposal in *CM* come online in the child's I-language at the same time. Perhaps I-languages initially allow only instructions for combining monadic concepts, and only add (limited) dyadicity later in the maturational process. If the latter, then we might ask, "Is this a *new* I-language?" More importantly, how could we tell?

On the assumption that I-languages have their lexical entries and composition rules essentially, we can adapt Pietroski's animals/molecules analogy, seeing a child's Slang as a succession of distinct I-languages—one for each day, week, month, or year. This entails that a child's Slang over the course of a year can be a set of significantly different I-languages. Though it is doubtless amenable to empirical inquiry, the question, "How many I-languages per day?" seems at best metaphysically awkward. Similarly, Pietroski holds that "dialectical variation … makes appeal to a single set of English expressions seem silly." Rightly so. But, again, the same considerations apply to the I-languages of speakers who are competent in many "dialects" or "languages" (as the benighted folk call them).

What are the individuation conditions on I-languages, in light of these considerations? In Pereplyotchik (2017: ch. 3), I argued that the answers to such questions are not much clearer in the case of I-languages than for the case of E-languages. I suspect that it will remain so for the duration of sustained inquiry in the decades to come.

4.2 *Philosophy has not failed cognitive science*

In a recent paper, provocatively titled "Why Philosophy Has Failed Cognitive Science," Brandom argues that analytic philosophy, exemplified in the work of Frege, has devoted a great deal of energy to clarifying the nature of logical and semantic notions, but that we've thus far failed to properly hand off the fruits of our heritage to researchers in cognitive science.[3] The present section is devoted to a survey of the claims that Brandom makes about this alleged failure. I'll argue that Pietroski's work provides a direct counterexample to several of these claims, but that Brandom is right to point out that many theorists, in-

[3] Recognizing that cognitive science is comprised of many fields, Brandom aims his criticism more directly at philosophers who work on topics in cognitive psychology, developmental psychology, animal psychology (esp. primatology), and artificial intelligence, *rather than* at those who study topics in neurophysiology, linguistics, perceptual psychology, learning theory, and memory. Admittedly, this is a strange way to cut up the terrain. In particular, for our purposes, it's not at all clear why philosophers of *linguistics* are not on Brandom's list of targets. But let's not dwell on this. If only for the sake of furthering our present inquiry, I'll include philosophers and language and linguistics in the list, making no exception for myself.

cluding Pietroski, hold commitments that Frege's insights should lead us to reject.

4.2.1 *Modes of inquiry, philosophical and scientific*

Brandom recounts the way in which modern approaches to logic and semantics began with Frege's *Begriffsschrift*, which furnished us with a new logic and new ways of thinking about meaning. Russell then showed us how to apply these ideas more generally in philosophy. But, while the ideas that Frege and Russell developed about logic and semantics were quite *general* in their import, later theorists to attempted, with variable success, to apply those general ideas to the specific case of natural language—i.e., the system of representation that normal human children acquire. This, no doubt, has to do our sheer familiarity with the only clear case of language use available—i.e., our own—coupled with the anthropocentrism that motivates any inquiry into language. The resulting confusion had the effect of blurring the lines between a general philosophical theory of language, on the one hand, and an empirical linguistic inquiry into the special case of human linguistic competence. That, Brandom maintains, is a mistake. On his view, the kind of inquiry that Pietroski is engaged in deals with the contingencies and the specifics of how humans acquired conceptual and linguistic abilities. Philosophy, by contrast, deals with the "normative" question of what *counts* as "doing the trick" for *any* creature.

Although I've followed Brandom in putting the point this way through the discussion so far, I must now register that this is not, in my view, the best way of saying what I think Brandom intends to say here. At any rate, it's not, by my lights, the point that he *should* be making at this juncture in the dialectic. For, one might legitimately wonder how "What counts as doing the trick?" gets to be a *normative* question—whether in the case of language or of anything else—rather than a straightforward question of fact. Presumably, "What counts as being a horse?" is not a normative question, for the simple reason that horses are a natural type of object, studied as such by zoologists.[4] Analogously, what counts as mastering a language may well be a matter of having acquired an I-language, in Pietroski's sense. As a friendly amendment to Brandom, then, I will address this worry on his behalf by re-iterating and fleshing out the proposal that I floated earlier in the discussion, regarding "levels of analysis".

As I've noted, what Brandom (and probably Lewis) has on offer seems to be a *high-level description* of a theoretically interesting kind of social practice—specifically, ("what counts as a") *language* game—

[4] A theorist sympathetic to Brandom might reply that the notion "counts as" is normative because it's a matter of what competences and abilities it is *appropriate* to ascribe to a creature. But, here again, a parallel move can be made in the case of horses, vis-à-vis the properties that are *correctly* ascribed to *them* (notably, the property of *being a horse*).

with all of the impressive social, practical, and cognitive benefits that make this kind of practice worthy of careful study. Correspondingly, I suspect Pietroski's proposal has its home in an inquiry pitched at a lower level of analysis. To flesh the picture out further, it will be useful to quickly rehearse a central tenet of the mainstream approach to "levels of analysis" in contemporary philosophy of science.

A view that posits multiple levels of theoretical analysis, whether in biology, computer science, or in the sciences overall, is not thereby committed to any particular story about how theoretical progress at any one level can, should, or must constrain theorizing at any other. True, the early proponents of a "levels of science" picture also attempted—unsuccessfully, as it turns out—to secure a "unity of science" thesis. But later thinkers, notably Fodor (1975), generously disabused us of these lofty goals. What we know now is that theoretical pressure can and often does "go both ways", with higher and lower levels informing one another in equal measure, and with equal authority. Lower levels, as such, are no longer seen as having an inherent epistemic privilege. This is a point that Chomsky, too, makes frequently.

Similarly, we can now appreciate the fact—poorly understood until fairly recently—that theories at different levels of inquiry are often to a large extent *independent variables*. A theorist who formulates a high-level analysis of some phenomenon typically assumes—often with good grounds—that the generalizations they discover at *that* level might be implemented in any number of ways by lower-level mechanisms. That's one half of the independence claim. The other half is best appreciated from the perspective of a theorist working at the (relatively) *lower* level of analysis. From *this* vantage point, the mechanisms, laws, generalizations, and/or principles that are discovered, however "abstract" they might seem, are assumed to be just one instance of an even *more* general phenomenon—a token of a potentially much larger type.

What I want to recommend is that we apply these general considerations from the philosophy of science to the concrete case of generative linguistics and normative inferentialism. Although it would be misleading to say that the subject matters that these two research programs seek to address are literally *orthogonal* to one another, the grain of truth in that bit of imagery is this: Brandom high-level account is, as such, indifferent to how lower-level mechanisms might operate in various token instances. Pietroski, on the other hand, is assaying the fine-structure of the *lower*-level mechanisms, but only in the special case of human languages. As such, while the results of his inquiry are *relevant* to Brandom's overall picture—indeed, they might pose devastating problems for Brandom (see below)—they function in practice *not* as substantive theoretical constraints, but as an account of a very special case (particularly to *us*!) of the kind of story that Brandom presupposes can be told for *any* creature to which his normative pragmatic account is applicable.

If this is so, then how do we make sense of the fact that Brandom, like Lewis, explicitly appeals to data from natural language in motivating his analyses of phenomena that are, at least in principle, specific to *our* way of doing things? If the theory is not intended to be a contribution to a "merely parochial" inquiry about *us*, then why use examples from *our* language—indeed, almost exclusively from English, in particular—in constructing and developing it? There are two complementary ways of answering this question. The first, already mentioned, is to point out that Brandom draws our attention to features of human languages (in practice, just English) *not* for the purpose of displaying empirical data that his account can explain, but, rather, to illustrate aspects of language that he believes have pragmatic or semantic analogues in languages beyond the human case. (By analogy, think of the Chomsky hierarchy.)

The second prong of the reply consists of highlighting the fact that Brandom has devoted much time and effort to arguing—ultimately persuasively, in my view—that many of the linguistic devices he treats in his work are actually *universal* features of language *as such*. Moreover, as many generative linguists have pointed out in discussions of Universal Grammar, language universals need not be categorical; they can, instead, take a conditional form, e.g., "any language that has feature F will also have property P," or "if a language can express content C, then it can also express content C*." Brandom (2008) works out a detailed typology of such relationships between logically possible languages, including those that differ either in respect of their general expressive power, or in respect of more specific semantic devices (e.g., deixis). He takes this to be a pragmatic-*cum*-semantic version of Chomsky's famous analysis of the syntactic hierarchies of expressive power.

The view that Brandom promotes throughout his discussions of this topic is that traditional philosophers of language, starting with Quine, directed their efforts at analyzing linguistic constructions that, by and large, shed light on quite general semantic phenomena—i.e., ones that we can hope to one day discover in other species (terrestrial or otherwise), or to build into our intelligent robots. Although such linguistic devices might seem, from the perspective of a modern-day linguist, to comprise a rather motley collection—why propositional attitude reports but not, say, ergative verbs?—the tie that binds them, according to Brandom, is one that we can best appreciate from the vantage point of a (high-level) normative inquiry into general pragmatics. The linguistic phenomena that Quine and others identified early on as being particularly germane to philosophy all have this in common: for each of them, there are good reasons to think that it's *not* just something we *happen* to find in distinctively human languages, but something that tells us about what a language *is*, irrespective of which creatures happen to use it or what subpersonal mechanisms they deploy in doing so.

4.2.2 *Frege's insights*

The lessons that Brandom believes philosophers have failed to pass on to their colleagues in the sciences pertain to four key distinctions, all due to Frege, between (i) labeling and describing, (ii) freestanding and embedded content, (iii) content and force, and finally (iv) simple vs. complex predicates.

The last of these, Brandom argues, opens up a semantic hierarchy that is no less important for cognitive scientists to be familiar with than the syntactic hierarchy that bears Chomsky's name. Taking this hierarchy into account in the context of empirical theorizing would help, he claims, to characterize the phylogenetic and ontogenetic development of linguistic and conceptual capacities. Such a characterization would move upward through what Brandom thinks of as "grades of conceptual content", including the propositional variety, the quantificational refinement, and ultimately the *relational* contents that Frege taught us to recognize.

We saw in 2 that Pietroski has a great deal to say about this. Indeed, the Fregean considerations that he surveys in the service of an avowedly naturalistic theory in cognitive science are precisely those that Brandom recommends to our attention (and *then* some). For Brandom, Frege's insight is that there are patterns in sentences that *cannot* be modeled as mere part-whole relations. For instance, although there is no expression that appears in "Herbie admires Jessica" and "Jessica admires Herbie" that doesn't *also* appear in "Herbie admires Herbie", the latter sentence exhibits an inferential pattern different from the other two—the pattern that we gesture at by employing notational distinctions between, e.g., admire(*x,y*) and admire(*x,x*), or by making explicit their inferential proprieties by embedding them inside of conditionals, as in (2) and (3).

(1) If someone admires anyone, then someone admires someone. (true)

(2) If someone admires anyone, then someone admires *themselves*. (false)

Thus, admire(*x,x*) expresses a kind of predicate that is not a *part* of a sentence, but an *aspect* of it, which we can recognize as an "inferential pattern" and *model* as an equivalence class of sentences. Frege's device of function-application is a way of capturing this idea. Functions are not, in general *parts* of their outputs. (The function *capitol-of(x)* yields Kiev when applied to Ukraine, but neither *capitol-of(x)* nor Ukraine are parts of Kiev.) This is why sentential connectives can be modeled with Venn diagrams, but complex predicates cannot. Even the simplest mathematics uses complex predicates—e.g., *natural number* or successor(*x, y*)—and Frege showed that, once you can build complex predicates, you can keep building endlessly more, in the manner we ran across in our discussion of Lewis's type-theoretic semantics (2).

As we've seen, Pietroski warns against taking for granted a creature's ability to construct concepts of unbounded adicities. But the warning is intended to apply *only* when doing natural-language semantics. For other purposes, Pietroski agrees that Frege's insights are of foundational and lasting importance. Moreover, the hypothesis that he develops posits thoughts that admit of a Fregean semantic treatment (perhaps even a truth-conditional one), but it requires these to first be converted, via Frege's process of concept invention, into the kinds of thoughts that are "legible", so to speak, to the human language faculty. While it's not clear what *independent* empirical evidence Pietroski might offer for positing psychological mechanisms that facilitate such a translation—I am aware of no obvious analogue in the case of other perceptual modules—what *is* clear is this: Brandom's contentions regarding Frege's distinction (iv), between simple and complex predicates, are rendered moot by the very existence of Pietroski's work, which presents an up-and-running empirical inquiry that is deeply informed by Frege's core contributions.

Matters are much less clear with regard to the other three distinctions that Frege was at pains to draw. Let's turn now to his distinction between labeling and describing.

4.2.3 *Sentences, predicates, and classification*

Brandom points out that old-school scholastic accounts of thought were rooted in a classificatory account of concepts—a relic of Aristotelian "forms". The medievals noticed that, once you have singular terms and classifications, you can build *up* to an account of truth, and then analyze good inference in terms of truth-preservation. Pietroski unabashedly endorses this strategy—in particular, the Aristotelian focus on classificatory concepts, which are central to his predicativism about New Mentalese (2).

This raises the question: What exactly *is* classification? How does a predicate get to perform its semantic function? Here is Pietroski's answer:

> …intuitively, a predicate classifies things, into those that meet a certain condition (e.g., being a rabbit) and those that do not. Anything that meets the condition satisfies the predicate, which applies to anything that meets the condition. We can invent perceptible predicates. Though for now, let's focus on predicative concepts, like instances of RABBIT. I assume that many animals have such mental predicates. … [A] predicate may apply to each of [several] things, or to nothing. But these are just special cases of classifying. …even if logically ideal predication is relational as opposed to classificatory, there seems to be a psychological distinction between relational and classificatory concepts, even if we speak of monadic/dyadic/*n*-adic predicates.

What I see, both here and throughout *CM*, are inter-definitions of semantic notions like "applies to," "classifies," "satisfies," and "meets conditions." Although Pietroski has made it clear that he is not trying to

"break out of the intentional circle," to use Quine's memorable phrase, the account he provides does not, to my mind, do much to illuminate the phenomenon in question. A diversity of labels allows us to conjure different clusters of theoretical intuitions. But none of these seems definite enough to make progress with.

Turn, then, to Brandom's answer, which has the virtue of laying out substantive proposals and refining them, arriving ultimately at one that meets various important desiderata. Like Pietroski, Brandom maintains that "classifying" is not the obtaining of some (super)natural relation between a concept and (a portion of) the actual world—let alone non-actual possible worlds. On his view, there are, instead, *acts* of classification—e.g., asserting "That's a rabbit," or tokening the corresponding perceptual thought ("Lo, it rabbiteth!"). We've already seen the details of Brandom's account of assertion, as well as his (subordinate) account of classification. Let's now approach the latter from a different direction, this time contrasting Brandom's view with extant rivals.

If asked, straight-out, "What is classification?," the knee-jerk response that most philosophers would offer is that classification is a matter of *differential responsiveness*. This is a start, but it leaves wide open the question of what vocabulary we're permitted to use in describing the objects, properties, and events to which a physical system might be differentially responsive. If we give ourselves free reign, then the notion becomes too cheap to do serious work; differentially responding to Italian and French operas would count as classifying them, regardless of how the trick was done. But, of course, one wants to know *how* that sort of thing happens, not just *that* it does. Unfortunately, pursuing the answer to this explanatory question by *restricting* our vocabulary to only naturalistically respectable terms quickly lands us with panpsychism—a bridge *much* too far. For, as Brandom points out, even a chunk or iron differentially responds to varying amounts of oxygen in its surroundings, e.g., by rusting.

Equally vacuous is the (unqualified) suggestion that we acquire predicative concepts, and hence classificatory powers, by performing a process of "abstraction" from either the intrinsic qualities of states of sentient awareness—as Hume, Russell, and Carnap all held at various points in their otherwise distinguished careers—or from the raw information supplied by sensory mechanisms, as naturalist like Neurath might have it. Without a detailed and well-motivated account of the operation of "abstraction", the acquisition of classificatory concepts has been labeled once more, but remains stubbornly unexplained.

To their credit, naturalists like Fodor and Dretske attempted to meet the problem head-on. Information-carrying states count as classificatory concepts, they argued, when they're embedded in suitably complex systems—ones that reliably keep track of their environment, learn, and behave flexibly, perhaps on account of their history of natu-

ral selection and innate resources. Burge (2010) adds to this list the requirement that the reliable tracking abilities must have the shape of perceptual *constancies*, not mere sensory registrations.

Brandom maintains that no such project can work, even in principle, precisely because it ignores Frege's conceptual distinction between mere *labeling* and full-blown *describing*. A case of labeling is one in which items are differentially sorted, but only extensionally, such that no specific inferential consequences can be drawn from the presence of the label. A magic wand might tell us that doorknobs, pet fish, and crumpled shirts are all and only the items that share the magic feature, F. But without knowing what F is, in intensional terms, we have no idea what, if anything, *follows* from the application of the label 'F'—i.e., what, in the fictional scenario, is semantically achieved by the activation of the wand. In order for this (or any other) physical signal to become more than a mere label, it must be inferentially articulated, in the sense that there have to be things that *follow from* something's being F, as well as things that can have an instance of 'F' in *their* inferential consequences.[5]

One of Frege's key lessons, then, is that inferential significance is central to conceptual content. Some concepts have only inferential conditions of application, not perceptual ones—either contingently, as with GENE, or necessarily, as with POLYNOMIAL or FRICTIONLESS PLANE. One can, of course, call things "concepts" even when they meet less stringent conditions. But, in that case, one should be sure to note the difference between differential responsiveness and inferential articulation. Moreover, these points hold irrespective of whether a differential-response capacity is innate or learned, and they apply just as much Boolean compounds of more basic units of differential responsiveness—i.e., *compound* labels.

While it's impossible to credit Brandom's claim that philosophers like Pietroski have taken insufficient notice of Frege's foundational insights, there is something to be said, I think, for his criticism on this

[5] Following Dummett's counsel, Brandom urges that we take into account *both* the circumstances *and* the consequences of applying a concept. For some nonsynonymous propositions, the antecedent circumstances coincide, but the inferential *consequences* serve to distinguish their contents. For instance, consider the contrast between "I will one day write a book about anarchism" and "I *foresee* that I will one day write a book about anarchism." The inferential antecedents ("circumstances") of these two claims might be the same, but the inferential consequences are different. This point applies even to observational concepts—e.g., MOVING or MOTION. A motion detector or a well-trained parrot that reliably emits the sound /*Moving*/ when there is, in fact, movement afoot (and not otherwise) does not thereby have the concepts in question. For although the circumstances of application are right, there are no inferential consequences to speak of in these cases. Brandom also makes the helpful observation that operators can serve to distinguish concepts that share both circumstances *and* consequences of application. For instance, the concepts HERE and WHERE-I-AM are shown to be distinct when interacting with the temporal operator 'always': "It's nice here/where I am" vs. "It is *always* nice here/where I am."

particular point. As we'll see below, Pietroski's views on classification don't seem to respect the distinction (whether it be Frege's or Brandom's) between labeling and describing. This has downstream consequences for Pietroski's view that really do seem to be out of step with Brandom's theoretical commitments.

The disagreement about classification is joined when Brandom asserts that *thinking about something*, as in "We're still thinking about his tax returns," is a matter of tokening *complete thoughts*—i.e., intentional states that can be expressed by linguistically competent creatures only in complete speech acts, which requires producing *complete sentences* (if only in the paradigmatic case). Pietroski, by contrast, follows the peculiar philosophical convention of using the phrase "thinking about" to denote a punctate event of conceptual classification. While he agrees with Brandom that having a thought requires tokening a "sentential concept", he also maintains that *all* concepts are "ways of thinking about things."

This is where Brandom would disagree, on account of his commitment to the effect that subsentential concepts are *not* complete thoughts. According to him, tokening such a concept cannot *by itself* constitute "thinking about something". To do that, subsentential concepts must (in some way) participate in a sentential one. So while sentential concepts are correctly described as "ways of thinking about things," Brandom follows Frege in viewing subsentential concepts as *aspects of* such ways. Thus, whereas Pietroski claims that "hearing 'Bessie' can… activate the denoter BESSIE, thereby leading [one] to think about Bessie in a singular way" (108), Brandom would deny that activating the denoting concept BESSIE can alone constitute thinking about Bessie—in *any* way—even *once*. This point about "denoters" applies also, *mutatis mutandis*, to predicative concepts.

Pietroski can, of course, *stipulate* that thinking about things doesn't require tokening complete thoughts. But it's difficult to see what could motivate such a move. Relying on brute introspection, one might fancy that singular reference has taken place with only one subsentential concept in play—e.g., "I'm quite certain that I was just thinking of tofu; not anything about it, specifically; just… *tofu*." However, such introspective judgments are known to be an extremely unreliable source of data, whether performed by naïve speakers or by theoreticians.[6] One might, more plausibly, appeal to the theory of *perception* developed by Burge (2010), according to which perceptual awareness involves the application of only *sub*sentential concepts, modeled on noun phrases. But this won't do, either. For, if judgments and classifications are all "sentence-sized", as Brandom argues, then even the *perceptual* mental attitude of *noticing* can't properly be treated as a case of applying just

[6] Distinct methodological troubles plague both of these two options, but they all strike me as insuperable and not worth discussing here. See Dennett (1991) for a primer.

one classificatory concept. Noticing rabbits involves *judging that* there are rabbits in the relevant spatiotemporal vicinity, and making such judgments requires deploying concepts other than the classificatory predicate, RABBIT(_)—e.g., the concept HERE(_).

With all this in mind, I think we should side with Brandom in saying that subsentential concepts *play a role* in acts of classification, where the latter are construed as either as public assertions or as inner endorsements of judgable contents. I see no reason to assume that tokening a subsentential concept is sufficient to carry off an act of classification. Nor is it obvious that classifying is a function of *all* concept application, as Pietroski believes. Does wondering whether Bessie exists really require *classifying* her? The latter question brings us face-to-face with the Fregean distinction between force and content, to which we now turn.

4.2.4 *Force and content*

Brandom draws our attention to an ambiguity that was long ago pointed out by Wilfrid Sellars—the so-called "-ing/-ed" ambiguity—which allows us to use words like "claim" and "thought" polysemously to describe speech acts and propositional attitudes in respect of their intentional contents, on the one hand, and in respect of their illocutionary force or "mental attitude type", on the other. With regard to the latter, Stephen Schiffer has popularized the imagery of different "boxes" in the mind—one that corresponds to the functional role of beliefs, another to that of desires, a third one for intentions, and so on. Pietroski likewise notes the distinction in the following passage from *CM*.

> One needs to be careful with the terminology, since words like 'thought' and 'concept' are polysemous with regard to symbols and contents; 'thought' and 'judgment' are also like 'assertion', which can be used to describe certain events that can be characterized in terms of contents. In speaking of a thought that Sadie is a horse, one might be talking about a mental episode, a mental sentence, or a content shared by various sentences.

Brandom goes on to argue that this distinction is not only useful for theorists, but that it also marks *a distinct level of conceptual sophistication*. Creatures who can tell the difference between the act of *asserting* and the content of what's *asserted* can be said to be aware, at least *implicitly*, of the force/content distinction. To make this awareness *explicit*, a creature can embed a sentence inside of a conditional, thereby stripping it of its force.[7]

Now, on the assumption that classification is, in fact, a kind of il-

[7] Brandom illustrates how conditionals can be used to distinguish those inferential consequences that derive from the *content* of what's said from those that derive from its *force*. Witness, for instance, the strikingly different inferential consequences of the sentences "p" and "I believe that p" when embedded as antecedents in conditionals: "If p then p" is obviously true for all values of 'p', but "If I believe that p, then p" is not foolishly arrogant for a mere mortal to assert, but also disastrously false in all known cases.

locutionary force, Brandom concludes that the assertion 'If F*a* then G*a*' cannot, in point of fact, be used to classify *a* as F, despite invoking both '*a*' and '*F*'.[8] This is another place where his views on the nature of classification come into conflict with Pietroski's. And, here again, I can think of no plausible way around it.

One might suggest, on Pietroski's behalf, that we seem to be able to simply *entertain a notion*—e.g., to contemplate "justice" or "the possibility of pigs flying"—without thereby committing ourselves to anything at all. This, Brandom points out, goes back to Descartes's view that one can first "entertain" an idea/proposition and then, *by an act of mental will*, either endorse or deny it, yielding either a committal judgment or a positive doubt. Pietroski's picture of concept-assembly likewise points in this direction. On that model, the process of assembly eventuates in the construction of a "polarized sentential concept", which is then shipped off to central cognition for endorsement, rejection, or further contemplation.

But this idea is at odds with Kant's equally compelling observation that concepts have contents only in virtue of their role in judgment. Pushing still further, Frege argued that entertaining propositions is a late-coming ability that involves a thinker embedding a proposition into the antecedent slot of a conditional—as in the following soliloquy: "What if *p*? Well, if *p* were the case, then *q* would also; but that would mean that neither *r* nor *s*…". If Frege's proposal is correct, then the ability to "entertain an idea" piggy-backs on two *prior* abilities—viz., to assert conditionals, and then to perform inferences that take *them* as premises or conclusions (e.g., hypothetical syllogisms).

Now, Pietroski agrees that the mental act of endorsement results in a committal judgment, which both he and Brandom take to be subject to normative evaluation—i.e., assessments of correctness, warrant, rational propriety and the like. But it's not clear how Pietroski's ⇑/⇓ operators for assembling polarized sentential concepts facilitate this act of endorsement. More generally, Pietroski's proposal seems to have little to offer in the way of a subpersonal about how *any* kind of force/attitude is superadded, so to speak, to polarized concepts, after the *Begriffsplans* get done assembling them.

4.3. *Predicates and singular terms*

We turn now to our very final topic, which concerns a foundational disagreement between Brandom and Pietroski on the nature of singular and predicative concepts. Recall that Pietroski's semantics for natural language is resolutely *predicativist*, in the sense that it recognizes no analogue of type-<e> expressions—intuitively, singular terms—i.e., no instruction for fetching singular concepts. Recall as well that he *does*

[8] Likewise, he warns against conflating denial and supposition—two kinds of force—with negation and conditionalization, which are semantic functions that directly participate in the content.

countenance the presence of such concepts in the human mind and that he recognizes the useful cognitive roles that such concepts play in thinking/reasoning. But this kind of cognition is couched in Olde Mentalese—the phylogenetically ancient representational format in which pre-linguistic thought was conducted, and which Pietroski thinks we still employ today, outside of language use.

Brandom develops a powerful argument to the effect that *any* language that fails to draw a distinction between predicates and singular terms is in principle barred from introducing basic logical operators—including both negation and the conditional. If this argument is successful, it would have no effect at all on Pietroski's claims about Olde Mentalese, which happily draws that distinction. But it would seem to present a rather major problem for Pietroski's main proposal about natural-language semantics, which has predicativism as one of its core commitments. So it behooves us, in surveying the points of discord between them, to focus on this foundational case, using it to draw out related points of contention about syntax.

4.3. *Brandom's argument*

"What are subsentential expressions?" and "Why are there any?" These are the two questions that Brandom raises in an essay of the same title (2001: ch. 4). In 1, we glimpsed the overall shape of his answer. Here, we'll reiterate the main points and look at some of the details. The reason for doing so is that this is the last—and arguably most challenging—of the issues that divide Brandom's normative inferentialism from the overall generative enterprise.

In the reconciliatory spirit of my overall project, I'll propose a possible strategy for ameliorating the dispute. But I should concede from the outset that this appears to be a particularly stubborn issue. This is frustrating, as the issue obviously cuts pretty deep. Having laid out the details of Brandom's difficult argument, I'll settle, in the end, for having *raised* the question—one that hasn't been discussed, to my knowledge, anywhere else in the literature—of how generative grammar might be (in)compatible with Brandom's substitutional approach to syntax.

4.3.1 *Details and a proof*

As noted earlier, Brandom agrees with Pietroski that discerning subsentential expressions is what makes it possible for us, both as theorists and as language users, to "project" proprieties governing the use of novel sentences. Once we've done this, we can then recombine subsentential items into new expressions, with meanings/contents that were previously inexpressible. Brandom recommends using the notion of substitution for this purpose, adapting Frege's insight that discerning meaningful subsentential expressions is a matter of treating sentences as *substitutional variants* of one another.

In spelling out the syntactic side of this technical notion, Brandom begins by identifying three "substitution-structural roles". These include the role of being an expression that is (i) substituted *in*, (ii) substituted *for*, and (iii) a substitutional *frame*. For instance, "David admires Herbie," is substituted *in* to yield a *substitutional variant*, such as "Herbie admires Herbie," where the expression 'David' has been substituted *for*. The residual substitutional *frame* is what's is common to the two substitutional variants—schematically, "*x* admired Herbie."

On the semantic side, a substitutional variant of a sentence will be defined in terms of the inferences that it enters into, as a premise or a conclusion. In keeping with his inferentialist project, Brandom develops the idea that the meaning of a subsentential expression consists in the materially correct *substitution inferences* involving that expression—i.e., inferences in which the conclusion is a substitutional variant of one of the premises. Thus, 'Herbie' has the meaning that it does partly in virtue of its role in a vast range of materially good inferences, including the single-premise inference from "Herbie barked" to "My dog barked" (as said by me).

With this in mind, Brandom notes that substitution inferences come in two flavors: symmetric and asymmetric. The above inference, from "Herbie barked" to "My dog barked", is symmetric, in the sense that it's materially good in *either* direction. Plainly, this trades on the identity between Herbie and my dog. This is more grist for Brandom's logical expressivist mill. He captures this observation by pointing out that identity is the logical notion that we use to *express*—i.e., make explicit—the *substitutional commitments* that are central to our notion of singular terms (and, relatedly, of the items they purport to denote). Contrast this with the inference from "Herbie runs" to "Herbie moves", which is materially good in only one direction, not in the other. That's because 'runs' is materially stronger, in respect of inferential consequences, than 'move'; the former licenses all of the inferences that the latter does, and then some. The distinction between symmetric and asymmetric inferential proprieties governing substitution inferences is, as we'll now see, the central aspect of Brandom's distinction between predicates and singular terms. Let's turn finally to his definitions of these two notions.

Each of the two definitions has a syntactic component and a semantic component. On the syntactic side, Brandom says that singular terms invariably play the substitution-structural roles of being substituted *for* (as well as *in*), whereas predicates invariably play the role of substitutional *frames*. On the semantic side, he points out that the substitution of singular terms is *always* governed by symmetric inferential proprieties, whereas predicates are *necessarily* governed by *at least some a*symmetric ones. For instance, 'Herbie' is a singular term partly in virtue of the fact that, if the substitution inference from "Herbie barked" to "My dog barked" is materially good, then so is its converse.

Crucially, the same does *not* hold for the substitution inference from "Herbie runs" to "Herbie moves", where the substitution of *predicates* is in play. That's, again, because 'runs' is inferentially stronger than 'moves'. This is an instance of something that Brandom goes on to argue is *constitutive* of predicates as a class—viz., that they *necessarily* enter into at least *some* asymmetric substitution-inferential relations with other predicates in the language.

Thus far, Brandom has supplied an answer only to his first question: what are singular terms (and, by extension, predicates)? To summarize, the answer is that singular terms play the *syntactic* roles of substituted *for*s and substituted *in*s, and the *semantic* role of entering *solely* into *symmetric* substitutional inferences. Predicates, by contrast, play the syntactic role of substitutional *frames* that *necessarily* enter into at least *some asymmetric* substitutional relations.

To ask why there are singular terms, then, is to ask the following question: Why do the syntactic and semantic substitutional roles line up as they do? This way of setting up the question allows us to generate a taxonomy of the logical possibilities, in terms of two binary parameters—syntax and semantics. We can thus imagine languages that instantiate the following four permutations.

i) Substituted *for* is *symmetric* and substitutional *frame* is *symmetric*.
ii) Substituted *for* is *asymmetric* and substitutional *frame* is *symmetric*.
iii) Substituted *for* is *asymmetric* and substitutional *frame* is *asymmetric*.
iv) Substituted *for* is *symmetric* and substitutional *frame* is *asymmetric*.

The option that's *actually* instantiated by singular terms and predicates is (iv). The question then becomes: What's "wrong" with the other options?

What rules out option (i), according to Brandom, is that, many of the substitution inferences that are to be codified and projected at the level of sentences by discerning subsentential expressions are *asymmetric*. No weakening inferences could be generated if *all* subsentential components were restricted *solely* to symmetric inferences. What the remaining options have in common is that they assign *asymmetric* inferential proprieties to expression-kinds that play the syntactic role of being substituted *for*. We can thus ask: what's wrong with *that* combination? The answer to this question is where things become technically challenging. Readers who feel like skipping ahead to the next section can take with them only the upshot of the proof: If a language fits the model of options (ii) or (iii), then it does not permit the introduction of conditional contents (contrary to fact, in our own case).

Brandom invites us to consider the generalizations that permit expressions with subsentential contents to determine the proprieties of a

productive and indefinitely flexible class of novel combinations. Asserting that Herbie is (identical with) my dog commits me to the propriety of *all* inferences of the form P(Herbie) → P(my dog). Similarly, for predicates; asserting that anything that runs thereby moves commits one to the propriety of all inferences of the form Runs(x) → Moves(x). This is why, when such content-constitutive and potentially asymmetric substitutional commitments made explicit, they take the form of *quantified conditionals*—another feather in the logical expressivist's cap.

Here, then, is Brandom's proof that options (ii) and (iii) in effect rob a language of its most basic logical notions—even ones as simple as negation and the conditional.

The pattern corresponding to the hypothetical *asymmetric* significance of "substituted *for*s" would replace identity claims with *inequalities*. Let "$t > t^*$" mean that P(t) → P(t^*) is in general a good inference, but *not* every frame, P, will make the converse inference, P(t^*) → P(t), materially good. Now, call a predicate Q an *inferential inverse* of a predicate P if, for all t and t^*, the following condition is satisfied.

Inferential Inverse $=_{df}$ if P(t) → P(t^*) holds, but P(t') → P(t) doesn't, then Q(t^*) → Q(t) holds and Q(t) → Q(t^*) doesn't

Thus, to answer the question of what's "wrong" with options (ii) and (iii), it suffices to show that if *every* sentential substitutional frame has an inverse, then there can be no *asymmetrically* significant substituted *for*s. The demonstration now proceeds by way of the following key lemma.

Lemma: In any language containing the expressive resources of elementary sentential logic, every predicate has an inferential inverse. Conditional and negating locutions are inferentially inverting; e.g., inferentially weakening the antecedent of a conditional inferentially strengthens the conditional. Thus, if condition the antecedent of *Inferential Inverse* holds, the then the consequent can be shown to hold as well.

Proof: Let Qα be defined as Pα r. It follows immediately that P(t*) →S(t*) entails P(t)→S(t), but P(t) → S(t) does not entail P(t')→S(t').

What this argument shows, if it shows anything at all (see below), is that conditional locutions are inferentially inverting precisely *because* they play the indispensable expressive role of making inferential relations explicit. (*Mutatis mutandis* for negation and other logical operators.) If this is right, then we can conclude, as Brandom does, that *any* language able to muster the expressive resources required for introducing basic sentential connectives will also draw a distinction between singular terms and predicates (as defined), assuming it has any subsentential structure at all. Conversely, any language that *forgoes* the term/predicate distinction is thereby severely castrated in its expres-

sive power—incapable in principle of introducing so much as a material conditional.

4.3.2 *Potential replies*

The foregoing argument was developed in Brandom's was presented in its canonical form in *MIE*. (See also the later and more condensed treatment in Brandom, 2001). In the decades since then, many theorists have marshalled a variety of technical objections against his line of reasoning. Some of these are based on straightforward confusions and can thus be defused without much concern (see *MIE*, ch. 6). Others might be more troublesome. Whatever the case about that, I want to ask what bearing this argument *would* have on Pietroski's position if it *were* successful.

As noted above, the argument appears to present a serious problem for Pietroski's commitment to predicativism, at least in the case of natural-language expressions and New Mentalese. (We can breathe easy about expressions of Olde Mentalese, which are in the clear.) How might Pietroski reply to this challenge? Closer to home, if this dispute can't be resolved, does that spell doom for my larger reconciliation project in the present essay?

One possibility is cut things off at the root by rejecting Brandom's substitutional approach to both syntax *and* semantics. Indeed, this is most obvious route for Pietroski to pursue, given his claim that we can't simply take for granted a creature's ability to "slice out" terms from sentences, so as to use them in combinatorically constructing an infinite hierarchy of semantic types, *a la* Frege or Lewis. Such a project, Pietroski argues, stipulates from the outset far more than it explains in the end. Suppose that he's right about this. Does that mean that his empirical results—assuming for present purposes that that's what they are—have literally *contravened* Brandom's strategy? Put another way, if generative linguistics is the correct approach to natural language, then are we barred from using Brandom's "substitutional scalpel" to identify subsentential structure, distinguish between singular terms and predicates, and carry off the inferentialist project at the subsentential level? I do not think so. Or, at any rate, I'm not convinced.

One excuse for my wavering on this point is that the considerations Brandom uses are so general—i.e., so totally independent of other details of the languages to which it applies—that it's hard to see which of them Pietroski is really in a position to deny. True, substitutional syntax smells a little too much like the old-school "discovery procedures" and "immediate constituent grammars" of benighted pre-Chomskyans (see Fodor, Bever, and Garrett, 1974 for a blistering refutation). But the methodological *similarities*, in my view, cut no ice. Nominalist discovery procedures, were, for all their shortcomings, *empirical hypotheses* about human languages. Otherwise, they wouldn't even *get* to be rendered false by straightforwardly empirical arguments. By contrast,

we've seen that Brandom's project, despite drawing on examples from English—again, *for illustrative purposes only*—is explicitly not pitched as an empirical inquiry into human language. So he can't be accused of attempting to resurrect that old idea.

Nor is it clear that Brandom's approach to the general project of delineating syntactic categories is incompatible with further elaborations by the kind of syntax that Chomsky supplies. In the only passage I've found where he mentions generative grammar and its transformation rules (Brandom, 1987), he makes precisely that suggestion:

> Recall … that Chomsky showed that one should not expect to generate the well-formed sentences of natural languages by concatenation, combination, or tree-structuring of any set of categories of this sort. To any such "phrase-structure grammar" will have to be added transformations of such combinatory structures. Categorial classifications are just the raw materials for grammar in this sense, and don't have anything to say about how one might proceed to the rest of the task of syntax once one has the categories. (165: fn. 2)

Lastly, we must consider how working syntacticians in the Chomskyan tradition go about identifying lexical items, morpho-syntactic categories, and other subsentential expressions. Needless to say, as a naturalistic, descriptive enterprise, the practice is subject to change under empirical pressures. Still, for the present day, the most common procedure is to employ what syntacticians call "constituency tests". For instance, the so-called 'do-so' test allows us to carve out phrases like the VP in (4) by reference to its behavior vis-à-vis (5), and then discern the more fine-grained syntactic units within that VP, by reference to (6) and (7).

(4) Jessica [$_{VP}$ swims quickly at the pool].

(5) Michael [does so], too.	'Does so' replaces 'swims quickly at the pool'.
(6) Aron often [does so] at the beach.	'Does so' replaces 'swims quickly'
(7) Hayes [does so] expertly in the tub.	'Does so' replaces 'swims'

Identifying such subsentential (and subphrasal) structure was crucial to the development of X-bar theory in the GB framework, and continues to guide syntactic theorizing in the Minimalist tradition. Of course, the empirical details are not so simple. Syntacticians employ a large battery of tests, not all of which agree with one another in every case. Moreover, any of the tests (including the do-so test) can be challenged on empirical grounds and, in many cases, rendered otiose by discoveries about the inner workings of *other* natural languages, only superficially unrelated. Mercifully, the gory details are not our topic here. For present purposes, the key point is this: When a syntactician employs constituency tests, she is appealing to precisely the kinds of substitutional relationships that Brandom's syntax rests on. To be sure, more complex considerations enter into the picture. For instance, how a syn-

tactic constituent can "move" (in grammars that allow movement, unlike say, HPSG) is a kind of data widely relied upon to determine syntactic category and constituency relations. But I see no reason why this too could not be cashed out in the language of substitutional syntax.

With this in mind, I'll end with the following speculation. If, in a remarkably distant possible world, Brandom *were* to go in for some empirical theorizing about natural languages, it's not clear to me that he would (or *should*) adopt anything other than generative grammar as the optimal theoretical framework within which to prosecute his inquiry. Certainly, he is well aware—how could he *not* be?!—of the theoretical need for Chomskyan grammars. On the rare occasion that he does mention these, he doesn't say anything that even hints at a disagreement. (Recall, "Chomsky *showed*…", my emphasis.) Moreover, his frequent invocations of the Chomsky hierarchy in discussions of computational procedures and the expressive power of various languages (e.g., Brandom, 2008: ch. 1) suggests no particular aversion to core generativist principles. To be sure, this isn't very much to go on. It by no means *shows* that Brandom's "substitutional syntax" is compatible with (any particular) generative grammar. But the consideration to the contrary seem likewise thin.

Conclusion

Attempting to integrate the theories developed by Brandom and Pietroski may strike some as an futile project, analogous to grafting, say, a squirrel onto a cow. One thinks to oneself, "Perhaps it *can* be done, but… why?!" In the foregoing pages, I've argued that this view of the matter constitutes a failure to appreciate the live opportunities for a fruitful merger. Such a merger is, like any large one, a daunting gamble. But, it seems undeniable, from where I sit, that both Brandom and Pietroski have furnished significant insights into the nature of something called "language"—a phenomenon that we should firmly resist regarding as unitary.

That having been said, it seems only natural to suppose that *combining* the two theories will yield a richer *overall* picture than either theory can provide on its own. This sort of thing doesn't always work out; not all teams of All-Stars are All-Star teams, after all. But even if the resulting view is not to one's liking, I find it frankly inconceivable that some such reconciliation project won't have to be effected *eventually*. Perhaps we aren't there yet; perhaps both generative linguistics and normative inferentialism must await more penetrating developments before their future descendants can be merged. (Or, again, maybe the AI people blow our minds with some new-fangled contraption next Tuesday. Who knows?) Whatever the case about that, I hope to have convinced the reader that there are, in fact, very few substantive *dis*agreements between the two approaches. What initially appear to be sharp contrasts turn out, on inspection, to be mostly benign differ-

ences of theoretical focus and explanatory ambition.

I'll close on a broadly sociological note. A mistaken commitment to the incompatibility of generative linguistics and normative inferentialism has had, I believe, negative consequences for both philosophy and linguistics. Specifically, there is, at present, little or no cross-talk between researchers working in these two traditions. Indeed, they seem to be as siloed off from one another as any two major research programs in "analytic" philosophy of language can be. If nothing else, by partially undermining the mistaken assumption of incompatibility, I hope to have gone at least some way toward rectifying the situation. My hope is that others will follow suit, attempting to forge still further connections between the two enterprises. Even if Pietroski and Brandom make for strange bedfellows, there is no question that they make for excellent guides. And, for better or worse, the terrain is largely uncharted. Let us press forward, then—as always, with optimism.

References

Brandom, R. B. 1987. "Singular Terms and Sentential Sign Designs." *Philosophical Topics* 15 (1): 145–167.

Brandom, R. B. 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge: Harvard University Press.

Brandom, R. B. 2001. *Articulating Reasons: An Introduction to Inferentialism*. Cambridge: Harvard University Press.

Brandom, R. B. 2008. *Between Saying and Doing: Towards an Analytic Pragmatism*. Cambridge: Harvard University Press.

Brandom, R. B. 2009. "Why Philosophy Has Failed Cognitive Science." In *Reason and Philosophy: Animating Ideas*. Cambridge: Harvard University Press.

Burge, T. 1973. "Reference and proper names." *Journal of Philosophy* 70: 425–439.

Burge, T. 2010. *Origins of Objectivity*. New York: Oxford University Press

Churchland, P. M. 1979. *Scientific Realism and the Plasticity of Mind*. Cambridge: Cambridge University Press.

Chomsky, N. 1986. *Knowledge of Language*. New York: Praeger.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge: MIT Press.

Chomsky, N. 2000. *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Chomsky, N. 1995. *Why Only Us: Language and Evolution*. Cambridge: MIT Press.

Collins, J. 2012. *The Unity of Linguistic Meaning*. Oxford: Oxford University Press.

Davidson, D. 1983. *Essays on Truth and Interpretation*. Oxford: Oxford University Press.

Dummett, M. 1978. *Truth and other Enigmas*. London: Duckworth.

Fodor, J. A. 1975. *The Language of Thought*. Oxford: Oxford University Press.

Fodor, J. A. 1979. "Three Reasons for not Deriving 'Kill' from 'Cause to Die'." *Linguistic Inquiry* 1: 429-438.

Fodor, J. A., Bever, T. and Garrett, M. 1974. *The Psychology of Language*, Cambridge: MIT Press.

Fodor, J. A. and Lepore, E. 1992. *Holism: A Shopper's Guide*. Oxford: Blackwell.

Fodor, J. A. and Pylyshyn, Z. W. 2015. *Minds Without Meanings: An Essay on the Content of Concepts*. Cambridge: MIT Press.

Grover, D. L., Camp, J. L., Belnap, N. D. 1975. "A Prosententialist Theory of Truth." *Philosophical Studies* 27 (2): 73–125.

Graff Fara, D. 2015. "Names are Predicates." *Philosophical Review* 124 (1): 59–117.

Grice, H. P. 1989. *Studies in the Way of Words*. Cambridge: Harvard University Press.

Harris, D. W. 2020. "We Talk to People, not Contexts" *Philosophical Studies* 177 (9): 2713–2733.

Heim, I. and Kratzer, A. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

Horwich, P. 1990. *Truth*. Oxford: Oxford University Press.

Kripke, S. 1980. *Naming and Necessity*. Cambridge: Harvard University Press.

Levinson, S. C. 1983. *Pragmatics*. Cambridge: Cambridge University Press

Lewis, D. K. 1969. *Convention*. Cambridge: Harvard University Press.

Lewis, D. K. 1970. "General Semantics." *Synthese* 22: 18–67.

Lewis, D. K. 1973. "Languages and Language." In K. Gunderson (ed.). *Minnesota Studies in the Philosophy of Science*. Vol. 7. Minneapolis: University of Minnesota Press.

Pereplyotchik, D. 2017. *Psychosyntax: The Nature of Grammar and Its Place in the Mind*. Springer International.

Pereplyotchik, D. 2019. "Whither Extensions." *Mind and Language* (Special Issue) 1–14.

Pereplyotchik, D. 2020. "Generative Linguistics Meets Normative Inferentialism: Part 1." *Croatian Journal of Philosophy* 20 (3): 311–352.

Pietroski, P. M. 2018. *Conjoining Meanings: Semantics Without Truth Values*. Oxford: Oxford University Press.

Price, H. 2011. *Naturalism Without Mirrors*. Oxford: Oxford University Press.

Price, H. 2013. *Expressivism, Pragmatism and Representationalism*. Cambridge: Cambridge University Press.

Quine, W. V. 1953. "On What There Is." In *From a Logical Point of View*. New York: Harper and Row.

Quine, W. V. 1960. *Word and Object*. Cambridge: MIT Press.

Quine, W. V. 1970/1980. *Philosophy of Logic, 2nd Edition.* Cambridge: Harvard University Press.

Rorty, R. 1979. *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.

Sellars, W. 1963. "Philosophy and the Scientific Image of Man." In R. Colodny (ed.). *Science, Perception, and Reality*. Humanities Press/Ridgeview: 35–78.

Stainton, R. J. 2006. *Words and Thoughts: Subsentences, Ellipsis, and the Philosophy of Language*. Oxford: Oxford University Press

Stalnaker, R. 1984. *Inquiry*. Cambridge: MIT Press.

Sperber, D. and Wilson, D. 1986/1995. *Relevance*: *Communication and Cognition*, *2nd Edition*. Oxford: Blackwell.

Tomasello, M. 2005. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.

Yang, C. 2006. *The Infinite Gift: How Children Learn and Unlearn the Languages of the World*. New York: Scribner.

Croatian Journal of Philosophy
Vol. XXI, No. 61, 2021
DOI

Book Reviews

Rafe McGregor, Narrative Justice, *Lanham: Rowman and Littlefield, 2018, 232 pp.*

Rafe McGregor's *Narrative Justice* presents an exciting development within the field of aesthetics. While offering an effort of great theoretical and practical significance—that of defending an account of narrative congitivism—its principal aim is to show that there is still a place for the humanities within education, and especially in practical use. McGregor's central thesis is that engagement with narratives has the capacity to improve the narrative sensibility of readers, which can lead to a decrease in criminal inhumanity—that is, in short, the methodology he refers to as narrative justice. If successful, his theory could provide the humanities and (especially) literature the means to resist neoliberal quantification, as it would demonstrate the value of humanities in terms of concrete individual and public gain. It is my impression that McGregor was to a very large degree successful in laying out a strong case for this endeavor and tackling possible criticism with great success.

It is important to mention the context within which this book is situated. Never before have the humanities faced such strong criticism from those trying to reduce most of the human activities to monetary profit. This is amplified especially by the COVID-19 outbreak which forces various governments to revisit their financial policies and to introduce measures like budget cuts to combat the potential threat of another financial recession. Within higher education and academia, the prospects are especially grim for the humanities, which are more likely to get their funding suspended than their colleagues working within the STEM area. Keeping this in mind, there has rarely been as much pressure on the humanities to defend themselves as there is now. An urgent need is present to contribute to the justification of the humanities if they are to have any future within higher education and academia. McGregor did just that by providing us with a theory that explains how engagements with narratives, both fictional and nonfictional, can have a valuable transformative impact on the audience and potentially reduce instances of harmful behavior.

However, McGregor is not alone in trying to situate the humanities, in this case the study of literature, within the context of neoliberal market/profit oriented society. Gayatri Chakravorti Spivak's theory of global aesthetic education and Sarah E. Worth's theory of narrative education stand out as the two most compelling contemporary alternatives to Mc-

Gregor's narrative justice thesis. While he agrees with most of what Spivak and Worth argue, McGregor finds shortcomings with both theories, rejecting them in favor of his own. However, a common core of the three authors is that they all argue for a conception of aesthetic education. The idea was developed by Friedrich Schiller, who argued that political harmony could only be achieved via the cultivation of aesthetic sensibility. However, McGregor puts his spin on this thesis via his account of the criminal inhumanity. He sees this as one of the crucial factors contributing to the increase or decrease of political harmony. The category of criminal inhumanity relates to serious crimes committed by a state or non-state actor against a civilian population, government or public for ideological reasons. With the emergence of various terrorist groups in the last three decades and subsequent terrorist attacks in cities around the world, there is a pressing need to combat terrorism in its modern form. The developments in the 21st century indicate that traditional hard power approaches that employ force as a tool of coercion have largely failed to pacify or eliminate terrorist hotspots in the Middle East—the armed conflicts in Afghanistan, Iraq, Syria, Yemen and countless others serve as a grim testament. This fact has given rise to soft power approaches that employ rhetoric as a tool of persuasion, rather than force. McGregor's thesis is situated firmly within this approach, drawing heavily on the insights delivered by the narrative turn in the sciences—most important of which is that narratives enable the acquisition of large portions of knowledge. In my view, part of the importance of the narrative justice thesis lies in the fact that it offers a theoretical framework that can be applied in practice to combat one of the most pressing global issues that our society has been struggling against for a prolonged period of time. As such, it has the potential to highlight one of the ways in which philosophy can indeed provide a positive impact on the world.

A central aspect of the book is the relation between narrative representation and ethical value. As McGregor argues, "The engagement with a narrative representation qua narrative representation is incomplete without ethical evaluation." (54). This line of thinking is designed to support the deflationary account of the ethical value of narrative representation. The idea behind this is that there is a necessary relation between narrative representation and ethical value, but not between narrative representation and moral value. Within the context of narrative justice, the term *ethical* refers to the opposition between something being moral and something being amoral, while the meaning of *moral* is linked to moral as opposed to immoral. McGregor claims that the essential value of narrative representation is restricted to the former, i.e., the ethical sphere. He maintains that the moral of the story may be virtuous, vicious or somewhere in between. Once established, the relation between narratives and ethics serves as a backbone for the further development of the main thesis of narrative justice.

For the thesis of narrative justice to be at all convincing, a theory of narrative cognitivism has to be established. If we argue that narrative representations provide us with a certain kind of knowledge, we have to provide a framework which enables knowledge to be transmitted. McGregor's aim thus is to defend the theory of narrative cognitivism, according to which narrative representations provide knowledge in virtue of their narrativity,

regardless of their truth value. He insists on two criteria for narrative cognitivism: the epistemic criterion (that the narrative representation provides knowledge) and the narrativity criterion (that the narrative representation provides knowledge in virtue of its narrativity). How is knowledge passed on to the reader? The idea is that the reader undergoes a kind of virtual experience while engaging with a narrative representation. To capture this mechanism, McGregor employs the term *phenomenological knowledge* (PK): "the realization of what a particular lived experience is like" (75). For the purposes of the narrative justice thesis the knowledge provided by exemplary narratives via *lucid phenomenological knowledge* (LPK) is of special interest. LPK is the "realisation of what a particular lived experience is like by means of the reproduction of a particular experience of a particular character for the audience who adopt the standard mode of engagement to the narrative representation" (76), i.e., that mode which is prescribed by the author. On McGregor's view, LPK meets the epistemic and narrativity criteria for narrative cognitivism.

To defend narrative cognitivism, McGregor considers a possible criticism issued at the narrativity criterion, namely lyric poetry. Poetry can be invoked by the critics so as to show that LPK is provided in virtue of the aesthetic properties, not in virtue of narrativity. This if of course a considerable burden for McGregor, since most of the discussion in literary aesthetics focuses on the question of aesthetic properties and their role in art's capacity to advance knowledge. McGregor undermines the objection by pointing to a narrative representation that, he claims, lacks aesthetic properties but provides us with LPK nonetheless. The example he uses is that of Morgan Spurlock's *Super Size Me*, a documentary that intends to show the negative health impacts of eating too much fast food. It lacks aesthetic properties, but, McGregor claims, is able to provide LPK by means of reproducing Spurlock's disgust towards fast food in the audience. While there are arguments that try to show the aesthetic properties of a documentary like *Super Size Me*, I think McGregor's defense in this instance is convincing. Thus, regardless of the problem of the nature of aesthetic properties, we can agree with McGregor that that is not the crucial issue when it comes to narratives providing knowledge.

At this point an argument for narrative justice can be laid out:

(a) The cultivation of narrative sensibility can develop ethical understanding
(b) Criminal inhumanity is a category of crime that is justified by ethical principles
(c) Theories have crime reduction potential in virtue of explanation (i.e., developing understanding of the causes of crime)
(d) Therefore, the cultivation of narrative sensibility has the potential to reduce criminal inhumanity.

In the last chapter of the book, McGregor demonstrates how narrative justice might work in practice. Specifically, he shows how a pair of exemplary narratives can be employed to reduce criminal inhumanity by undermining extremist recruitment strategies. The application of the narrative justice thesis should result in the realisation that the Muslim fundamentalist and white supremacist ceonceptions of victimhood are identical and therefore

the victim master narrative on which both of the ideologies are based comes out to be false. One line of reasoning employed in this chapter states that if the groups are opposed to one another (which they obviously are), they should tell different stories. As it turns out, both groups are telling the same story, which then leaves only two options: either the groups are not actually in opposition, or both stories are false. Judging by their ideology both groups are clearly in opposition which leads us to the conclusion that both of their stories must be false. By performing a comparative analysis of two texts coming from groups that belong to each extremist ideology, Ajit Maan reveals that they present two instantiations of the same master narrative:

(i) In the beginning, our people lived in utopia
(ii) Then others arrived and took over
(iii) This brings us to the present, where we have two choices
(iv) We can either do nothing, in which case the situation will remain as it is now, or we can expel these others and restore the utopia in such a way that it is never threatened again

We can see that both texts share the same narrative form and the same conception of victimhood; McGregor calls this core conception *deliverance*. This line of reasoning enables us to conclude that the concepts of White genocide and Crusader are two instantiations of deliverance.

This example of undermining inhumanity and others that feature in the book supports the claim that the comparative analysis of documentary and fictional narratives has the potential to reduce criminal inhumanity. As such, the method with which narrative justice could be applied in practice comes in the form of the careful selection, analysis and comparison of documentary and fictional narratives for the purposes of disclosure, demystification and deconstruction. McGregor built a strong theoretical framework that supports the thesis that criminal inhumanity can be reduced by narrative sensibility. I think that narrative justice can withstand most of the theoretical objections raised against it, but it might encounter difficulties in practical application. Before any theory of aesthetic education could hope to be implemented in practice or influence the creation of new state policies, it should be backed up by empirical data. However, McGregor is not optimistic on the possibility of such empirical support. He spends a considerable amount of time analyzing the current empirical research regarding the impact of different forms of art, including violent video games and pornography, on people, only to conclude that "the attempt to demonstrate a link between literature and empathy is thus far inconclusive" (120). McGregor thus agrees with Spivak that most if not all of the benefits of literature are unverifiable (126), but claims that there is a necessary relation between exemplary narratives and lucid phenomenological knowledge. However, this is a practical problem for any theory of aesthetic education: while the theory itself may be sound, its application, which should be its principal aim, faces obstacles that greatly hinder its implementation. Thus, the lack of supporting empirical evidence is one of McGregor's most pressing challenges, but it is important to note that this is not a criticism of narrative

justice but rather a recognition of one of the difficulties the humanities will have to face. Surely we can expect interdisciplinary research of philosophy, literature, criminology, psychology and other fields of knowledge to be important with respect to meeting this difficulty and the pointers that McGregor provides in this book are much needed!

Narrative justice proves to be a project with the potential to answer two very important questions that have been raised in contemporary academia: (a) the question of how should one go about the defense of the humanities against the pressure coming from neoliberal quantification and (b) the question of which (if any) approach could be employed in order to combat the rising threat of extremist violence and terrorism that is present throughout the modern world. The central thesis of narrative justice that the cultivation of narrative sensibility could reduce criminal inhumanity has the potential to address both of these questions and therefore it has to be taken seriously. I think that McGregor demonstrates that the thesis of narrative justice holds against most of the criticism on the theoretical level. I identified one of the problems with it, namely difficulties involved in its application, and I argued that any theory belonging to the context of aesthetic education will face those issues. Thus, my claim should not serve as an objection to narrative justice per se, but rather highlight the enormous amount of work and empirical research that will have to be done before any step towards practical application can be undertaken. I conclude that narrative justice proves to be an important development within the wider context of philosophy and the humanities. I strongly recommend this book to everyone interested in ways in which narratives and ethics interact, and in the wider social context within which we run the risk of inhumanity.

MATIJA RAJTER
University of Rijeka, Rijeka, Croatia

Tiziana Andina, Petar Bojanić (eds.), Institutions in Action. The Nature and the Role of Institutions in the Real World*, New York: Springer 2020, 150 pp.*

Philosophy is a humbling profession. Even after 20 years of "dabbling" in it, I am often confronted with the Socratic "I know that I know nothing", or at least its more moderate version "I wish that I knew more". The same sentiment in me was also provoked by this book, as I went into the task of reading it with the preconception that I am familiar with what institutions are, only to be confronted with the different aspects on the subject that have once more proved that I was wrong in my hubris.

But, let us focus on the book. *Institutions in Action* by editors Tiziana Andina and Petar Bojanić is a part of the Springer's series "Studies in the Philosophy of Sociality," which main aim is to provide important contributions to the rapidly developing field of social ontology. As social ontology is the field of study that investigates corelation of nature with the social world, its focus is on various entities that appear in the world as a product

of social interaction. As such, social ontology is especially interested in the analysis of social groups, but other investigated entities also include social classes, genders, races, language and of course, institutions.

In general, social ontology could be regarded as a branch of metaphysics, as its research is mainly concerned with the nature of entities. However, it is also hard to strictly demarcate the social ontology's field of research, as its investigation of social entities oftentimes overlaps with one of social sciences. The difference lies in the approach, which in social ontology is more detailed, it asks the "big questions" about the nature of our social world. Furthermore, social ontology as a term has been widely used in recent times, but its topic has been a subject of philosophical inquiry since ancient times.

As editors in the preface of the book claim, their aim was to collect contributions from distinguished scholars from various theoretical and methodological orientations, in order to investigate the theory related to the nature of institutions, their identity and normativity, with the overall goal of bringing the debate back to the centre of social ontology. To do that, the book's main aim is to answer a simple question—"What are institutions?" Of course, this question begets new questions, such as, how can we define institutions, and can we arrive on a single agreeable definition? Of course, *Institutions in Action* does not stop there, but also tries to answer related questions whether institutions could be considered collective subjects, or do we have to understand them differently, as a class of specific objects with particular properties? And if the institutions are understood as a class of specific objects, what kind of entity they would be and what properties would define them? And finally, what kind of relationship is there between the institutions and the singular subjects and what and how many types of relationships can there be between institutions?

As it is evidenced by the questions, ontological inquiry takes precedence in this book to other areas of philosophy, such as moral and political philosophy, that also commonly deal with the institutions. But it does not mean that it evades the questions relevant to that fields, as the question of the relationship of institutions with the individuals is intermittently invoked throughout the book. For instance, the question whether institutions are collective subjects or are they different from the sum of individuals that compose them has repercussions on political and ethical debates, as prolonged existence over time (which is one of defining properties of institutions) enhances the trust between the institutions and citizens. And the trust is necessary condition for the proper functioning of institutions, that is, protection of individuals through the application of justice.

Finally, "Institutions in Action" also deals with the relation of institutions and elements that have normative basis, such as norms, laws, and contracts. This relation, in turn, provokes questions such as "can we reasonably understand normativity as the foundation of institutions?" And if the answer is positive, should we think of normativity as external or internal to the institutions?

First of the essays in this interdisciplinary collection is "Social Corporations as Social Institutions" by one of the leading authors in the field of social ontology, the Finnish philosopher Raimo Tuomela. In it he argues that "social corporations" (e.g., school, hospital, mail service) ideally function as

a kind of "extended" social institution. Such an extended institution, in turn depends on more basic institutions like language, money, and property. It also provides services by their host group for its target group. Tuomela argues that ideally properly functioning social institutions should be based on "full-blown we-thinking," as well as the collective acceptance of a fact or property as institutional. Moreover, the creation of institutions ideally requires we-thinking in which group members together create an institution by the act of declaration (or at least by an agreement). Tuomela's account gives evidence for the claim that the social corporation that in general is "by the people for the people" functions as a social institution in the extended sense. In more general terms, a social corporation exists in a host community (or state) and involves a normative organization with positions – thus statuses and powers – for the individuals.

Francesco Guala and Frank Hindriks in their "Institutions and Functions" tackle the general questions on the ontology of institutions by using a functionalist approach. They identify the types of institutions by their function or the coordination problems they solve. Their idea is that this functionalist approach provides some insights into the limits of reform, or the extent to which institutions like marriage, property, or democracy can be modified without turning them into entities of a different kind.

Snježana Prijić Samaržija in her "Epistemic Virtues of Institutions" considers how through the use of our everyday speech we attribute moral values not just to individuals, but also to institutions. By doing so we assume that institutions are bearers of epistemic features for which they can be blamed or applauded. In the essay, she inquires how can we responsibly ascribe epistemic features to institutions and asses the meaning of attributing virtues and vices to institutions. Finally, she offers the preliminary position that an epistemically virtuous institution is an institution which applies scientific knowledge (knowledge about what is real) in everyday life (knowledge about how to act) in order to improve the quality of overall life for all (knowledge about what is good or bad for citizens).

In the title of his essay "What Is an Act of Engagement? Between the Social, Collegial and Institutional Protocols" Petar Bojanić already presents us his main research question. In it he tries to describe this special kind of social act and to determine the function these acts have in relation between various agents. He also tries to define their significance in the transformation of a group into an institution or higher order entity. His premise is that there are acts whose aim is to engage all others, since they refer to all of us together, and in so doing reduce negative (social) "acts", as well as various asocial behaviours within a group or institution. In that manner, he considers that engaged acts could alternatively also belong to a kind of institutional act, since they introduce certain adjustments to the institution, by changing or modifying its rules, and increasing its consistency and efficiency.

Next essay "Play It by Trust: Social Trust, Political Institutions and Leisure" by Nebojša Zelič tries to explore the connection between institutional arrangement and social trust. First, he discusses why trust is important for democratic institutions, but also why it is an indispensable element of the welfare state institutions. Second, he outlines the various understandings of trust, in which he gives special emphasis on social trust. Third, he

describes how institutions contribute to generating and sustaining social trust. And finally, he asks us to seriously consider impacts of institutions on our leisure time, which could be understood as a catalyst of social inclusion, which is, in turn, the precondition of social trust.

In "Individual Morality and the Morality of Institutions" Thomas M. Scanlon discusses the relationship between moral philosophy and political philosophy, that is, he discusses the morality of institutions and the content of moral standards governing the relations between individuals and institutions. This issue leads him to conclusions on individual morality, which deals with standards applicable to individuals. On the basis of a contractualist conception of individual morality and a conception of the morality of institutions that follows Rawls' theory of justice, his article addresses the question of the foundations of the obligation to comply with institution-defined standards that are directed towards individuals. Finally, Scanlon focuses on the difficulty of rationalizing that obligation in the case of unjust institutions.

"States and Transgenerational Actions" by Tiziana Andina aims to analyse the metaphysical structure of transgenerational actions and to show that they are a necessary condition for the existence of social reality. She suggests that, in order to understand the structure of social reality, it is not enough to consider it as a complex architecture with a certain structure at a certain point in time. It is also important to analyse its basic elements, paying particular attention to its components that change over time. Social reality needs transgenerational actions in order to exist and endure. In other words, there can be no social reality, organized as we know it, without transgenerational actions.

"From Capital to Documediality" by Maurizio Ferraris is a complex text bursting with ideas in which he highlights what he calls a "documedia revolution". He starts with Documediality which indicates the allegiance between the constitutive power of documents and the mobilizing power of media. He outlines this great transformation, showing how we, as humankind, moved from Capital to Documediality, passing through Mediality. It is his opinion that from Documediality, one must start today to understand the ongoing social transformations as well as human nature, leaving Capital to historians.

On the other hand, "The Basis of European Cooperation" by Jonathan Wolff tries to provide a philosophical explanation for a scheme of cooperation that is European Union. He argues that to understand the purpose of the EU it is essential to debate about the underlying understanding of cooperation. This, in turn, will have implications on issues such as immigration policy and political union. Unfortunately, this debate has not taken place to the extent it is required, and as Wolff's analysis suggests the Brexit is one of the consequences.

The final text in this collection, "Ways of Compromise-Building in a World of Institutions" by Emmanuel Picavet represents an exploration of the institutional challenges during the proposed implementation of general principles. He pays special attention to principles such as rights, which receive a recognised moral interpretation. In the article, he argues that the implementation process is basically associated with power allocation issues,

which cannot be reduced to lateral, technical problems, but on the contrary, they help shape the institutional effectivity and collective significance of principles.

To conclude, the book does deliver on its promise on providing an interdisciplinary view on topic of institutions. Still, this is not a book that will give you simple and clear answers if you just want to familiarize yourself with the subject. On the other hand, if you want to further your knowledge, or you just need a fresh perspective in your research on the topic of institutions, the articles by eminent scholars compiled in *Institutions in Action* are a great place to start.

MARIN BEROŠ

United Nations Development Programme, Human Development Report 2020. The Next Frontier Human Development and the Anthropocene, *director and lead author Pedro Conceição, New York: United Nations, 412 pp.*

The series *Human Development Index* published under the auspices of the UN has been for three decades of its existence an invaluable source both of data and of analyses of global development, its achievements and problems. Developed by the economist Mahbub Ul Haq, it has been inspired and supported by the work of the Nobel laureate Amartya Sen; indeed, the present volume begins by Sen's tribute to Ul Haq, who passed away in 1998. The series offers invaluable materials to theoreticians interested in global development, and, relevantly for our reader, to philosophers working on global issues and on perspectives of cosmopolitan approach to ethics and politics. The present volume is particularly interesting in this respect, since it deals with the most burning global issues. The issues, collected here under the general title of "Anthropocene" encompass ecological themes, prominently climate change and pollution, and health issues, made recently prominent by the new pandemic. Thus, the book sketches the new playground on which the cosmopolitan debate is taking place, quite different from the traditional one that has inspired pioneering contemporary authors, like Peter Singer, some decades ago.

The Part One entitled "Renewing human development for the Anthropocene" develops this new diagnosis, and the last, Part Three offers the quantitative data. However, the Part One concludes by pointing to the direction in which the problems diagnosed could be solved, and the road to solution is the topic of the central part, Part Two, entitled "Acting for change". "We are entering the Anthropocene: the age of humans" the section begins. And goes on by claiming: "For the first time in our history the most serious and immediate, even existential, risks are human made and unfolding at planetary scale." (20–21; the numbers in brackets stand for pages). The "Anthropocene" the word was coined and the concept introduced by Eugene Stormer in the 1980s and popularized by Paul Crutzen in the 2000s (see Frank Biermann 2018 Global Governance in the "Anthropocene," in Chris Brown and Robyn Eckersley (eds.) *The Oxford Handbook of International Political Theory*, Oxford University Press, 467–479).

Let me point to some factual characterization offered in the book. It lists climate change and then mentions the phenomena of biogeochemical cycles disrupted (especially with nitrogen and phosphorus), ocean acidification, land-use change and biodiversity loss (49). This goes together with abundance of new materials of pure anthropogenic origin (aluminium, concrete, plastics), presence of radionuclides linked to atmospheric nuclear weapons testing, altering the diversity, distribution, abundance and interactions of life on Earth. Then come increasing species extinction rates, habitat losses, overharvesting, and growth of invasive species (49).

The next phenomenon characterizing the new playground, are the pandemics."A new normal is coming. Covid-19 is the tip of the spear", we are warned (4). All these factors are then taken to construct new indices of human development. The basic idea is to identify the amount of "planetary pressure", the ecological pressure that a given country place on the planet Earth (235). The two main components of the pressure are the CO_2 emissions and the "material footprint per capita". For the first component, emissions are attributed to the country in which they physically occur. The second one concerns biomass, fossil fuels, metal ores and nonmetal ores. It is calculated as raw material equivalent of imports plus domestic extraction minus raw material equivalents of exports. Material footprint per capita describes the average material use for final demand.

This yield a new top lists. The present index has Ireland at the first place, then Switzerland, followed by Great Britain. Hong-Kong's position is unclear, but then we have for certain Denmark, Germany, and then Sweden on the safe sixth place. Croatia, the home country of this journal, has made a huge progress, 19 places upwards and is now No. 43 on the list. On the other hand, Australia falls 72 places in the ranking, while the United States and Canada fall 45 and 40 places respectively, due to their disproportionate impact on natural resources. China drops 16 places from its current ranking of 85.

So much about the Index. But let me note that the descriptive part does not stay with the diagnosis of the problems; Part One, characteristically, concludes, with possible future directions. Let me quote a typical passage:

> Social cohesion and mitigating inequalities are enablers—not just prerequisites—for human development
>
> It was frequently emphasized that a reconceptualization of human development that addresses cohesion across and within society—relations between countries or across generations and relations with nonhuman natures and ecologies—is threatened by a grossly unequal world and by the narratives, technologies and processes that perpetuate inequalities. Social cohesion requires vertical and horizontal trust within societies while respecting diversity of beliefs and worldviews. Enhancing social cohesion, mitigating inequalities and restoring the value of social and socionatural relations require the inclusion of multiple voices and perspectives. (113)

The main part of the book is dedicated to such possible solutions of the problems listed in Part One. Its first chapter, the fourth in the book as the whole, bears characteristic title "Empowering people, unleashing transformation". The next is entitled "Shaping incentives to navigate the future", and the last in Part Two "Building nature-based human development". The reference to the beginning of HDI points to the continuity of motivation: The transformational changes required to ease planetary pressures and redress social imbalances call for another reorientation of goals and choices like the one that the HDI encouraged 30 years ago (227).

The volume connects ecological issues with political matters in a clear and quite radical way: How can we use our power to expand human freedoms while easing planetary pressures? The Report argues that we can do so by enhancing equity, fostering innovation and instilling a sense of stewardship of the planet. It stresses the need to empower the disadvantage groups: "Inequalities in empowerment today are at the root of environmental problems, many threatening the wellbeing of future generations. Important for a better tomorrow is to empower disadvantaged groups and actors today" (77).

They encompass disadvantaged groups in general and developing countries in general. Even more demandingly, one should empower "everyone through knowledge, change in norms and stewardship of nature", as well as various majorities in various communities, as well as local communities themselves. The list also refers to indigenous peoples. Finally, the political changes have to be global: "Changing incentives to preserve biodiversity is difficult given the complexity of the fabric of life. A key challenge is that biodiversity remains undervalued in current markets, despite the increasing appreciation of its contributions to people" (172).

The chapter, most importantly, notes the following: "For climate change and biodiversity loss, individual actions and even national actions will not do enough to ease planetary pressures" (173).

Philosophers working on the prospects of cosmopolitanism can appeal to these suggestions to argue that the present-day statist arrangements, even supported by UN and similar institutions' optimism, are dramatically insufficient to deal with problems on the new playground. We shall return to this point in a moment.

So much about the main claims and main proposals of the book, with apologies for all the topics there was no space-time to mention here. Let me now pass to the discussion. First, a linguistic-conceptual question. The book features "Anthropocene" already in its title. But is the term right? Is it really the work of the human species, of all of us? Or is it rather the work of the richest, a small minority of humanity?

Interestingly, the positive answer is suggested at least by one author in the book, Gaia Vince, in her chapter. Vance is the author of a very thorough recent (2014) book on the topic *Adventures in the Anthropocene: A Journey to the Heart of the Planet We Made* (Chatto and Windus). She notes that "the richest people in the world are doing the most to damage the environment that we all rely on for clean air, water, food and other resources" (2014: 121). And she adds that "they experience few consequences and the least danger from this environmental damage. The richest 10 percent of the

world's population are responsible for half of carbon emissions, while the poorest 50 percent are responsible for just 10 percent." Further, she tells the reader the wealthiest people contribute less socially, paying in the least to the collective pot. Here is her illustration: "In relatively equal Scandinavia the richest 0.01 percent illegally evade 25 percent of the taxes they owe, far higher than the average evasion rate of 2.8 percent.17 In the United States the richest 400 families pay a lower effective tax rate than any other income group. An estimated $9–$36 trillion is stored in tax havens around the world" (121). She notes optimistically that how poor people get rich will strongly shape the Anthropocene.

Given that there is quite a lot of agreement in the literature that the relevant phenomena are primarily the work of the rich minority, I disagree that "Anthropocene" is the right term. The work is not done by us, the non-wealthy majority of the mankind, and the has been misnamed. I propose to call it rather "Plutocene", like in "plutocracy" from Greek: "ploutos", wealth.

Let us pass to more principled topics. *The Report* is published under the auspices of the UN, and is proposing measures that would fit an advance but still statist global system. However, the reader might often feel that the tasks specified, as part of what it would be "Acting for change", as the title of the Part Two of the volume suggests, central for the global therapy of humanity are very hard to be fulfilled within the firmly statist system.

Start from the claim we quoted above that "Human development is possible only within planetary boundaries". The economic development, authors claim, should be "reinterweaved" with environment (113). And the enablers for this are "social cohesion and mitigating inequalities" (113). The take-home message we quoted tells that enhancing social cohesion, mitigating inequalities and restoring the value of social and socionatural relations is urgently needed, and that this task requires the inclusion of multiple voices and perspectives. This sounds like an appeal to very tightly knit world, and it seems quite impossible to achieve the required social cohesion, enriched with "vertical and horizontal trust" within a firmly statist global society: such social cohesion seems to require political and trust-supported connections at various levels, bringing together "multiple voices and perspectives", which normally disappear in a purely statist perspective.

Similarly with the demands of equity, we quoted above. For the authors "procedural equity relates to how decisions are being made in reference to institutions, governance and participation (64). And they claim that "…inequalities in empowerment today are at the root of environmental problems, many threatening the wellbeing of future generations. Important for a better tomorrow is to empower disadvantaged groups and actors today (77).

The same requirements re-appear when we look at slightly less general considerations. Take as the example the list of "actors to be empowered", we just quoted. Who could effectively empower developing countries? With the statist system, it will be the biggest and strongest states that will decide; who could force Trump's US or Putin's Russia to empower developing countries? Who would accept to empower indigenous peoples and their local communities all over the world? The present day Bolsonaro's Brazil is a good example of the impotence of the international community in the face of a determined cynical despot, supported by the white urban majority within

his country. Who can empower "everyone through knowledge, change in norms and stewardship of nature"?

This kinds of tasks demand a differently organized global community, say a supra-nationally controlled loose federation or confederation, where the sovereignty of states is dispersed regionally and globally. And only a strong democratic control could guarantee the literal empowerment of everyone "through knowledge, change in norms and stewardship of nature", as the requirement goes.

Similarly with biodiversity. Remember our quote from Chapter Five: " For climate change and biodiversity loss, individual actions and even national actions will not do enough to ease planetary pressures (173). And let me add that, quoting Elinor Ostrom, the authors note that what is needed is a polycentric approach to the problem of climate change and talk of "enhancing international and multi-actor collective action." Indeed, Ostrom herself, in the 2009 paper referred to, speaks of "multiple scales and decisionmaking units" ("A Polycentric Approach for Coping with Climate Change, The World Bank Development Economics," Office of the Senior Vice President and Chief Economist, October 2009, 39).

To conclude, I would suggest that the best way to read all these suggestions is in the direction of a less centralized, more multi-leveled spheres of influence. However, this goes again beyond the statist global model. We need a polycentric global model, and this seems to point beyond the UN-style statist system. It is encouraging that the official publisher, the UN, is publishing and promoting a book that pretty much suggests both its principled impotence and the existence and the appeals of a much more cosmopolitan solution to the problems from the present-day playground.

NENAD MIŠČEVIĆ
University of Maribor, Maribor, Slovenia

Croatian Journal of Philosophy is published three times a year.
It publishes original scientific papers in the field of philosophy.

Croatian Journal of Philosophy is indexed in
The Philosopher's Index, PhilPapers, Scopus, ERIH PLUS
and in *Arts & Humanities Citation Index* (*Web of Science*).

Payment may be made by bank transfer

SWIFT PBZGHR2X
IBAN HR4723400091100096268

Instructions for Contributors

All submissions should be sent to the e-mail: cjp@ifzg.hr. Submissions must be in English and formatted to be double-spaced with suitably wide margins, an A4 page size, and automatic page numbering.

Articles are normally no longer than 8,000 words of main text including bibliography. The Journal will consider longer papers, but, once these limits are exceeded, authors should bear in mind the editorial policy that the acceptance bar raises with increasing length.

Manuscripts should be compiled in the following order: cover page; title; abstract (not exceeding 200 words); keywords (3 to 6); main text; appendices (as appropriate); references.

All the authors of a paper should include their full names, affiliations, postal addresses, telephone and fax numbers and email addresses on the cover page of the manuscript. If a paper is co-written, one author should be identified as the Corresponding Author. The cover page must be submitted as a separate document. All submitted manuscripts must be prepared for blind review, with revealing acknowledgements and self-identifying references removed.

Sources are cited in the text by the author's last name, the publication date of the work cited, and a page number if needed, e.g. (Barber 2007: 324). Full details appear in the reference list in which the year of publication appears immediately after the author's name:

Barber, A. 2007. "Linguistic Structure and the Brain." *Croatian Journal of Philosophy* 21 (7): 317–341.

Williamson, T. 2013. *Identity and Discrimination*. Oxford: Wiley-Blackwell.

The publication of a manuscript in the Croatian Journal of Philosophy is expected to follow standards of ethical behavior for all parties involved in the publishing process: authors, editors, and reviewers. The journal follows the principles of the Committee on Publication Ethics (https://publicationethics.org/resources/flowcharts).

ISSN 1333-1108
9 771333 110001