

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

## Different Kinds of Fusion Experiences

### This is the author's manuscript

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/1722558> since 2020-01-12T19:52:37Z

*Published version:*

DOI:10.1007/s13164-019-00456-7

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

## Different Kinds of Fusion Experiences

### *Abstract*

Some people have stressed that there is a close analogy between meaning experiences, i.e., experiences as of understanding concerning linguistic expressions, and seeing-in experiences, i.e., pictorial experiences of discerning a certain item – what a certain picture presents, viz. the picture's subject – in another item – the picture's vehicle, the picture's physical basis. Both can be seen as fusion experiences, in the *minimal* sense that they are experiential wholes made up of different aspects. Actually, two important similarities between such experiences may lead one to think that they are experiences of the same type. In this paper, however, I will try to show that notwithstanding such similarities, these experiences are typologically different. For there are two dissimilarities between such experiences that are more relevant on this typological concern than their similarities. Indeed, unlike meaning experiences, seeing-in experiences are first of all recognitional experiences of a sort that makes them perceptual experiences (though *sui generis*) as well. Moreover, again unlike meaning experiences, seeing-in experiences are fusion experiences in the *substantial* sense that the experiential whole is more than the sum of its experiential parts taken in isolation, for when such parts figure in that whole, they interact with each other. In a nutshell, unlike meaning experiences, seeing-in experiences are *proper* fusion experiences.

### *1. Preliminaries*

Consider first an experience as of understanding, or a *meaning* experience concerning a linguistic expression (Strawson 1994:7). One has this experience when, in hearing (seeing) certain sounds (marks) even morphosyntactically structured, one hears (sees) those sounds (marks) as having a certain meaning (from now on, I will basically focus on the auditory version of the experience). Second, consider the experience of *seeing-in*, the pictorially relevant experience of discerning a certain 3D scene, i.e., what the picture presents viz. the

picture's *subject*,<sup>1</sup> in another (typically 2D) object, the picture's *vehicle*, i.e., the material or better physical basis of the picture (Wollheim 1980, 1987, 1998, 2003a,b). Now, both experiences can be taken to be *fusion* experiences, in the *minimal* sense that they are experiential wholes made up of different aspects, a lower-level and a higher-level one (Husserl 2002, Briscoe 2018, Wollheim 1987). In the former case, such aspects are i) the auditory perception of certain morphosyntactically organized sounds and ii) the proper meaning component of the meaning experience. In the latter case, such aspects are i) the (visual) perception of the picture's vehicle and ii) the experience of the picture's subject. Wollheim (1998) labels i) and ii) respectively the *configurational fold* (CF) and the *recognitional fold* (RF) of seeing-in.

Ordinary language seemingly supports the idea that there is a close relationship between those experiences. Indeed, we say not only that one sees a subject (a pictorial content) *in* a picture's vehicle, but also that one hears a thought (a linguistic content) *in* the expression(s) one hears (McDowell 1998). Yet notoriously, the vagaries of ordinary language provide no reliable evidence for substantial matters, in this case for the idea that there is a *typological* commonality between such experiences. Nevertheless, whether or not they have explicitly defended it, influential philosophers have flirted with this idea. On the one hand, in (1991, 2009<sup>4</sup>) Wittgenstein thought that the two experiences should be conceptually co-investigated. For, he said, both experiences are *aspectual* experiences that someone who were blind to the relevant aspect (either a meaning or a pictorial aspect) could not entertain. Wittgenstein's use of the expression "blind" is not metaphorical at all. For Wittgenstein thought that in meaning experiences we can *sensuously* experience meanings, just as in seeing-in experiences we so experience the picture's subject. On the other hand, in the first edition of his (1980), Wollheim himself was analogously seduced by the analogy he found between meaning experiences and pictorial experiences. For, he stressed, just as we do not infer the meaning of an expression from the sounds that we hear, but we hear a sort of unseparated whole, we do not infer the perceptual properties of the picture's subject from the perceptual properties of the picture's vehicle, but we again have an integrated experience.

Now, there certainly are two important similarities between such experiences. By following those philosophers' suggestions, one may think that such similarities are enough in order to justify the claim that, in their being fusion experiences in the above minimal sense,

---

<sup>1</sup> In this paper, I will assume that what the picture presents coincides with its subject, taken as what the picture is about. Actually, this assumption is rightly controversial. See fn. 9.

meaning experiences and seeing-in experiences are typologically identical. In the end, we will see that such a claim is not justified. Yet, the hypothesis of their typological identity is worth pursuing. Here are the similarities in question. In both cases, first, the higher-level aspect existentially depends on the lower-level aspect; this dependence, second, is a generic one.

Yet as I just said, such similarities notwithstanding, in this paper I want to hold that a meaning experience and a seeing-in experience are not experiences of the same kind. The plan of this paper is indeed to show how these experiences ultimately differ, so that they amount to different kinds of experiences. There are indeed between such experiences two dissimilarities that are more relevant for their typological, actually different, nature, than the above similarities. First, unlike a meaning experience, a seeing-in experience is a *recognitional* experience. In particular, how seeing-in is recognitional contributes to explaining why it is also a *perceptual* experience, though *sui generis*, while a meaning experience is just a *cognitive* experience that is merely grounded in the perceptual experience of an expression in its morphosyntactical form. Second, in its perceptual way of being a recognitional experience, unlike a meaning experience a seeing-in experience is a fusion experience in a *substantial* sense originally pointed out by Stumpf (1890), i.e., a *proper* fusion experience; namely, an experience in which the whole is more than the sum of its experiential parts taken in isolation, for when such parts figure in that whole, they interact with each other. Indeed, unlike a meaning experience, a seeing-in experience actually fulfils the requirements for being an experience of that sort, given how its CF determines its RF.

The architecture of this paper is as follows. In Section 2, I outline the two similarities between meaning experiences and seeing-in experiences. In Section 3-4 I try to show why, these similarities notwithstanding, the two experiences are typologically different, because of their more relevant dissimilarities. In Section 4 in particular, I conclude that, unlike meaning experiences, seeing-in experiences are proper fusion experiences.

## 2. The similarities between meaning experiences and seeing-in experiences

As I hinted before, there are important similarities between meaning experiences and seeing-in experiences that a defender of the typological commonality between meaning

experiences and seeing-in experiences may appeal to. Two such similarities deserve to be pointed out. True enough, I believe that in the end these similarities do not manage to give proof of any typological commonality between meaning experiences and seeing-in experiences. Yet, they are worth being explored, for I agree that they indeed exist.

First of all, both experiences are such that their higher-level aspect depends on their lower-level aspect, where this dependence is first of all an *existential* one.

Let me consider meaning experiences first. As regards such experiences, their higher-level aspect, i.e., the proper meaning experience concerning a certain expression, would not exist if their lower-level aspect, i.e., the perceptual experience of that expression (in its morphosyntactic structure) did not exist as well.

At first blush, one might object that, as regards meaning experiences, there is no such dependence, by stressing that straightforwardly *meaning* something by means of a certain expression does not depend on hearing that expression. For when one straightforwardly means something, one does not need to hear the expression one uses in order to mean that something. For example, when one utters “Watch out!” in order to warn someone of danger, one does not need to hear that sentence in order to so meaning it.

In response, a defender of the typological commonality might remark that even straightforwardly meaning something presupposes entertaining some sort of quasi-perceptual auditory (or visual) imagination in the inner speech that accompanies that act of meaning; namely, either a form of mental imagery or even a form of imaginative perception of the expression’s morphosyntactic organization (Carruthers-Veillet 2011, Prinz 2011, Tye-Wright 2011).

Granted, the aforesaid opponent might not accept this response, by denying, in an antiprivatistic Wittgensteinian fashion as it were, any role to inner speech in straightforwardly meaning something.

Yet the defender of the typological commonality may further reply that, even if there were something like a mere straightforwardly meaning something, this would constitute no counterexample to the claim that in meaning experiences, their higher-level aspect, the proper meaning experience, depends on their lower-level aspect, the auditory perception of an expression. For the above objection merely affects the *production* phase of language, but not its *comprehension* phase. Now, meaning experiences viz. as of *understanding* experiences just concern the second phase. Comprehending is precisely affected by the

above dependence of the proper meaning higher-level aspect of the meaning experience on the lower-level aspect of the auditory perception of the relevant expression.

Thus all in all, as regards meaning experiences, we may take the existential dependence of the higher-level aspect, the proper meaning experience, on the lower-level aspect, the auditory experience of the relevant expression, as firmly established.

Let me now pass to seeing-in experiences. It clearly seems that the same can be legitimately said also as far as seeing-in experiences are concerned. The RF of a seeing-in experience would not exist if the CF of that experience did not exist as well (Hopkins 2008).<sup>2</sup>

Indeed, this dependence sounds indisputable. In seeing-in experiences, we would not discern a subject in a picture if we did not already see the picture's vehicle. Indeed, we can discern that subject by virtue of seeing the vehicle (as we will see later, by virtue of seeing it in a particular 3D way). If one erased the colors and forms constituting the basic perceptual properties of the vehicle, we could no longer discern any subject in it.

So, let me grant the first kind of similarity between meaning experiences and seeing-in experiences, concerning the existential dependence of their higher-level aspect on their lower-level aspect. Moreover, there is a second kind of similarity between meaning experiences and seeing-in experiences. This kind regards the *particular* form of dependence that in both experiences the higher-level aspect has on the lower-level aspect: namely, a *generic* dependence. Let us see this point in detail.

On the one hand, in its higher-level aspect, i.e., the proper meaning experience, a meaning experience depends *generically* on a perceptual experience of the relevant expression, since it depends on *some* such experience or other. Not only can one perceive the same expression in different modalities (typically, hearing and seeing) while still experiencing the same meaning, but it is also the case that this higher-level aspect of a meaning experience remains the same, even if different perceptual experiences in the same modality of different yet synonymous expressions are mobilized, either in the same or in different languages that one understands. As regards an *intralinguistic* case, consider these two admittedly synonymous sentences:

---

<sup>2</sup> Husserl (2006) holds a similar idea as regards the relationship in pictorial experience between its experiential component that concerns what he calls the *image-object* – again, what the picture presents (which for him differs from the picture's subject, taken as what the picture is about) – and the experiential component concerning the picture's vehicle (cf. Eldridge 2018).

- (1) Furze has yellow flowers
- (2) Gorse has yellow flowers.

Whoever is competent in English hears the same meaning in them (obviously by virtue of the fact that “furze” and “gorse” are (admittedly) synonymous in English), although the perceptual experiences of (1) and (2) are different, for “furze” and “gorse” sound differently. As regards an *interlinguistic* case, take the French and the Italian sentences that allegedly translate both (1) and (2):

- (3) Les ajoncs ont des fleurs jaunes
- (4) I ginestroni hanno fiori gialli.

Provided that one is competent both in French and in Italian, one’s meaning experience does not change, even if of course (3) and (4) respectively sound different from both (1) and (2). Thus, as regards a meaning experience, the generic form of dependence of the higher-level aspect on the lower-level aspect is firmly established.

On the other hand, as regards a seeing-in experience, one might immediately deny that things are the same. For one might *prima facie* think that the dependence of the RF on the CF of that experience is *specific*, not generic: if there is a different CF, then there is also a different RF. To take an artistic example, consider what happens when a painter – Marcel Duchamp – slightly alters the colors and shapes that are instantiated in the vehicle of Leonardo’s *La Gioconda*. As a result, in entertaining the seeing-in experience corresponding to such alteration, one entertains a different CF that goes with a different RF, in which one sees the woman with a moustache constituting what Duchamp’s *L.H.O.O.Q.* presents.

This *prima facie* thought notwithstanding, even as far as seeing-in experiences are concerned the dependence of the RF on the CF is (at least up to a certain extent) *generic*. Indeed, a certain range of variation in the CF is allowed that does not induce a corresponding variation in the RF of a seeing-in experience. There are various examples showing that this is the case. In a first and hardly controversial group of examples, consider a black-and-white photo getting older and becoming sepia. Although in re-experiencing it

the CF of the relevant seeing-in experience obviously changes, it still presents the same scene, so that the RF of that experience remains the same. In the same vein, consider a TV set transmitting a colored shot that all of a sudden, perhaps because of a blackout, turns into a black-and-white one. Various people would agree that the shot still presents the same scene (cf. e.g. Wittgenstein 1977); hence again, a change in the relevant CF is not matched by a change in the relevant RF.<sup>3</sup> Let us now pass on to a second, actually more controversial, group of examples, to be traced back to Wollheim 2003b. In Parmigianino's *Madonna with the Long Neck*, says Wollheim, we do not see a woman whose neck is abnormally elongated, but just a normal (possibly Renaissance) woman. Thus, if Parmigianino had painted a picture that differed from his masterpiece only as regards the length of the lines tracing a neck, in it one would have still seen the same scene, thus entertaining the same seeing-in experience with the same RF even though its CF was different. Ditto for another example Wollheim provides there; namely, Matisse's *Green Stripe*. We see in it not an alien woman with a green stripe on her face, but a normal woman, the very same woman we would see if Matisse had not painted green the relevant part of his painting.

Admittedly, as I hinted at before, this range of variations in the CF of a seeing-in experience has limits. It is not the case, for example, that in *any* picture of the 'Madonna and Child'-type, one has the very same seeing-in experience. In order for this generic dependence to hold, first of all, the scene presented in the RF must remain precisely the same, in order for different CFs to contribute to the same seeing-in experience. Moreover and perhaps less obviously, other criteria, on top of preservation of scene identity, must be provided to this concern, as we shall now see.

Concerning these criteria, the first of the above two groups of examples can be easily accounted for as cases of generic dependence of the RF on the CF by appealing to the fact that, different colors in the picture's vehicle (the photograph, the TV set...) notwithstanding, in the CF of the relevant seeing-in experience *the same 3D-like organization of the vehicle* is preserved. To anticipate an example I shall discuss later, if an 'aspect dawning' picture like the picture of a Dalmatian were such that its white spots turned into pink ones, the very same Dalmatian-like silhouette would appear in it. Thus, we would still see an ordinary Dalmatian dog in it (as if we were seeing it through pink glasses). Theoretically speaking, this would also occur if the alteration of a shape, rather than that of a color, were at stake.

---

<sup>3</sup> Some people oppose this claim, by holding that what one sees in the altered shot is rather a black-and-white subject matter: Husserl (2006), Nanay (2016, 2018). For some reasons against this opposition, cf. AUTHOR 2; see also below.



As I said, the second group of examples of is more controversial. Indeed, it is unclear whether appealing to identity of 3-D like organization of the vehicle may be enough in the Parmigianino case (and in the Matisse case as well) in order to show that the dependence of the RF on the CF of the relevant seeing-in experience is generic. For it is unclear whether the shape alteration that would be at stake if one painted a shorter shape in the area of the vehicle where a neck is seen would preserve the same seeing-in experience. (Hopkins 1998 holds that in the Parmigianino case (and in the Matisse case as well) one should draw a distinction between what one actually sees in the picture – a woman with an elongated neck – and what the picture presents, its figurative or pictorial content – a woman with a normal neck; see also Brown 2010.) One may put the perplexity this way. As anyone knows, Amedeo Modigliani is universally known for having depicted fascinating yet neck-elongated human beings. Certainly enough, if while observing most of Modigliani's paintings one did not see in them human bodies whose neck was elongated, one would miss the point of the pictures. Thus, if there were a nonstandard painting by Modigliani that differed from a standard Modigliani merely because it did not contain the relevant elongated shapes, observing it would hardly preserve the same seeing-in experience. Why should the same considerations not hold as regards the Parmigianino painting (or for the Matisse painting for that matter)?

Yet even in those cases, the generic dependence of the RF on the CF of the relevant seeing-in experience is well confirmed. Indeed, Wollheim 1980 may help us in finding another criterion for distinguishing between the Parmigianino-like cases and the Modigliani-like case, including the Matisse example (see also Bantinaki 2018) that assesses that dependence of the RF on the CF is generic. According to this further criterion, we may stick to the idea that, unlike the Modigliani-like cases, in the Parmigianino-like cases what is seen in a picture whose vehicle contains some deformity is still a nondeformed scene. In looking at pictures there is, first of all, not just seeing-in, but *correct* seeing-in; namely, what one *must* see in a picture. Moreover, correct seeing-in is fixed by *authorial intentions*, i.e., by what the picture's creator wants us to see (and it is possible for us to see) in the picture (the proviso is important: if a creator wanted us to see something in something else that was impossible for us to see, as when a baby makes a scribble and says "That's Dad", there would be no such 'Dad'-seeing-in experience). Thus, unlike the Modigliani cases, in Parmigianino's *Madonna* we see a normal, not an alien neck-elongated, woman. For, unlike Modigliani, Parmigianino wanted us to see this (which was possible for us to see). Once this further criterion of *intentionally correct seeing-in* is adopted, the appeal to an (albeit limited) generic dependence of the RF on the CF is reinforced. For by virtue of this criterion,

one may say that if the Parmigianino painting were suitably altered (i.e., normal neck-like shapes occurred in it), one would still see the same scene in it. Ditto for the Matisse painting. Certainly, this would not be the case as regards a Modigliani painting that did not contain the relevant elongated shapes. Yet, as regards any Modigliani painting, it would be the case as regards other intended features (for example, if a Modigliani painting were slightly altered in its background colour, it would still allow one to see the same scene in it).

Thus, all the above examples – both the humble ones coming from photos and TV sets and the superb ones coming from painterly pictures – show that the dependence of the RF on the CF of a seeing-in experience is generic. Concerning the very same examples, this point can be restated also by appealing to another sort of consideration. As Wollheim (2003a) insisted, seeing-in is *cognitively penetrated*, notably in a *strong* sense (Macpherson 2012, 2015), as far as its RF is concerned: namely, the content of that fold is *constituted* by concepts of the items that the picture presents. Thus, we do not see black-and-white aliens in pre-color TV shots because we expect them to present ordinary beings. If we did not know that black-and-white pictures are meant to depict normally colored flesh-and-blood individuals, we would quite likely see black-and-white items in them. Likewise, we see a normal woman in Parmigianino's *Madonna* because we know what sort of thing is to be seen in it, given that we know the kind of painterly genre it belongs to. Or to appeal to another example put forward by Hagberg 2016, once we know that Franz Auerbach's *Sketch from Titian's Bacchus and Ariadne* is a contemporary attempt at mimicking Titian's *Bacchus and Ariadne*, we see normally human-like beings in it, not outlandish sketchy items. Yet on the contrary, if we did not know that this sketch were meant to be a sketch from that well-established masterpiece, in it we would quite likely see outlandish sketchy items.

All in all, therefore, one may plausibly say that not only in meaning, but also in seeing-in experiences, the higher-level aspect depends generically on the lower-level one. Thus, not only the first similarity between meaning experiences and seeing-in experiences – existential dependence of the higher-level aspect on the lower-level aspect – but also the second similarity between such experiences – generic dependence of the higher-level aspect on the lower-level aspect – is confirmed. But are these similarities enough to prove that such experiences are experiences of the same kind? As we will see in the next Section, the answer to this question is negative.

### 3. The dissimilarities between meaning experiences and seeing-in experiences: 1

In this Section I want to argue that, the above similarities notwithstanding, there are at least two fundamental differences between meaning experiences and seeing-in experiences that show that such experiences are typologically different.

Here is the first difference: unlike meaning experiences, in seeing-in experiences one can read off the higher-level aspect (the RF) from the lower-level aspect (the CF).

Let us consider meaning experiences first. However complex a meaning experience is, its lower-level aspect – the auditory perception of the relevant expression (where this expression is morphosyntactically individuated) – yields no clue to get its higher-level aspect – the proper meaning experience. Granted, there is a phenomenal shift between hearing an expression and having an overall meaning experience affecting it. For one can hear an expression, not only in its lower-level perceptual features (sounds) but also in its higher-level perceptual features (morphosyntactic organization), and still experience no meaning in it, at least if by “meaning” one means *lexical*, if not even *truthconditional*, meaning.<sup>4</sup> Yet, once one entertains a meaning experience, the previous auditory perception of the expression along with its morphosyntactic organization remains the same as in hearing the expression so individuated in isolation. Simply, the overall meaning experience is the outcome of the fact that a proper meaning aspect is juxtaposed to that auditory perception.

To see that this is the case, consider first of all an instance of *satiation* involving a meaningless expression. Take the meaningless expression “bly” and repeat it until you no longer hear that morphological structure, but rather the different morphological structure corresponding to hearing the different meaningless expression “libe”. What you have had until that perceptual switch is the auditory perception of a mere meaningless expression along with its own morphological structure. After that switch, you have a different auditory perception of another mere meaningless expression along with its own morphological structure. Suppose now that “bly” acquires a particular meaning; subsequently, you will experience that expression with that meaning. Clearly enough, there is a phenomenal shift between the first experience, the auditory perception of the mere meaningless expression

---

<sup>4</sup> I do not want to deny that grasping a certain morphosyntactic structure in a (meaningless) expression goes along with grasping a *structural* form of meaning for it, as all philosophers of language endorsing a Chomskyan attitude as to meaning matters maintained. By hearing the meaningless *Jabberwocky*-like sentence “the mome raths the borogrove”, someone competent in English certainly grasps that something did something to something else. Yet she grasps neither the lexical meaning of its constituent expressions nor their truthconditional contribution, for there is none.

“bly” in its morphological structure, and the second experience, the overall meaning experience affecting that expression. Yet that shift is constituted by merely adding an higher-level meaning aspect to the auditory perception of that expression in the above structure, thus ending up having an overall meaning experience. *Mutatis mutandis*, the same holds for “libe” once it acquires a particular meaning: the experience of that meaning is an aspect merely juxtaposed to the previous auditory perception of that expression in its own morphological structure.

Moreover, consider to the same purpose a more complicated situation. Take first the meaningless *Jabberwocky*-like:

(5) The slithy toves gyred the Jabberwock in the wabe

as heard in the morphosyntactic organization that allows it to ‘say’ that in their being in the wabe, the slithy toves gyred the Jabberwock, instead of ‘saying’ the structurally alternative ‘reading’ that the slithy toves gyred in the wabe the Jabberwock. There is definitely a phenomenal change from hearing (5) in its first structural ‘reading’ to hearing it in its second structural ‘reading’. At this point, suppose that each of the lexical entries in (5) acquires a certain simple meaning, so that in its first morphosyntactic organization, (5) compositionally acquires a certain complex meaning. Thus second, you experience (5) in that complex meaning. Clearly enough, there is a phenomenal shift between the first experience, the auditory perception of a mere meaningless sentence in a particular morphosyntactic structure, and the second experience, the meaning experience with that sentence in that structure. Yet that shift consists in merely adding an higher-level meaning aspect to that auditory perception of the meaningless sentence in its morphosyntactic organization, thus ending up having an overall meaning experience: again, the experience of that meaning is an aspect of that overall experience that is juxtaposed to the previous auditory perception of that expression in its own morphological structure. *Mutatis mutandis*, the same holds of (5) taken in its second morphosyntactical organization when each of its lexical entries acquires a certain simple meaning.<sup>5</sup>

---

<sup>5</sup> One might remark that this argument is simply based on introspective evidence. In one sense, this remark is correct. The argument is based on two cases of phenomenal contrast – the contrast between the auditory experience of a meaningless yet morphosyntactically articulated expression and the experience of that

Against the idea that in meaning experiences the higher-level aspect is simply juxtaposed to the lower-level aspect, one might rebut that knowing the meaning of an expression changes the experience of that very expression *per se*. Knowing that meaning enables one to no longer hear mere noises, but to hear a linguistic expression (O'Callaghan 2011). Thus, when it figures as the lower-level aspect of a meaning experience, the auditory perception of an expression is no longer the same as when it occurs in isolation.

But this is an improper way of putting things. For if such knowledge enables one to hear more than mere noises, this means that as far as the relevant expression is concerned it also enables one to hear also *its morphosyntactic structure*, which was not previously grasped in hearing those noises. The resulting experience, however, is *not* a meaning experience; it is just the auditory perception of the expression *along with* its morphosyntactic structure that is prompted by that knowledge of the meaning. *This* – the auditory perception of both the sounds and their morphological structure – *is the kind of experience that remains the same* once it figures as the lower-level aspect of a meaning experience. For a proper meaning aspect concerning that expression is *merely added* to that auditory perception. In a nutshell, although knowledge of meaning may prompt one to auditorily single out a morphosyntactically individuated expression, from so hearing that expression one cannot *read off* that meaning. As it can be shown also by the fact that one may so hear that expression even if one does not *know* its meaning, but simply one (erroneously) believes that it has (another) one.

Analogously, it might seem that certain sounds feel familiar once they are knowingly given a certain meaning, so that they are no longer heard as they were before, in isolation.

But again, even this way of putting things is improper. Independently of whether it is prompted by any meaning knowledge, the feeling of familiarity may also occur just on the basis of morphosyntactically organized sounds alone, even if such sounds are not experienced as having a meaning. This happens precisely with the meaningless (5), which is quite familiar to many of us. Indeed, however prompted, the feeling of familiarity essentially

---

expression as endowed with a certain meaning, and the contrast between an expression in a certain morphosyntactic articulation and that expression in another such articulation – and sometimes at least, phenomenal contrast cases rely on introspection (Chudnoff 2015:56). Yet in another sense, the remark is incorrect. For the second case of phenomenal contrast simply serves to show that, as regards a meaning experience, its first aspect is identical to the auditory experience of an expression in a certain morphosyntactic articulation *taken in isolation*; in other words, that such an experience is unaltered by figuring as an aspect of a meaning experience. As we will immediately see, this is not the case with the perception of a picture's vehicle *qua* mere physical object and the perception of it *qua* pictorial vehicle, i.e., the CF of a seeing-in experience.

just concerns the mere auditory experience of the expression in its morphosyntactic structure. *This* experience remains the same once a higher-level meaning aspect is *merely added* to it, in order to have an overall meaning experience. So again, one cannot read off that meaning from hearing that morphosyntactically individuated expression.

By contrast, in a seeing-in experience things stand differently. Indeed, as regards this experience one can *read off* from seeing its CF what its RF amounts to. For the RF is not juxtaposed to the CF, since the CF amounts to a *substantial modification* of the experience of the picture's vehicle *taken in isolation*.

Against this idea, one might immediately rebut that one can see a picture's vehicle *per se*, i.e., like any other physical object, thereby grasping both its lower-level perceptual properties and its 2D higher-level organizational properties. However, by so seeing the vehicle, one sees no scene in it, nor has any idea of what scene there is to be seen. In this respect, there is definitely a phenomenal shift between perceiving that vehicle *per se* and seeing a scene in it. Incidentally, this is the kind of shift Gombrich (1960) appealed to; namely, a shift between seeing a vehicle *qua* physical object among others and seeing it *qua* picture.

Fair enough: up to this point, I agree, no relevant difference occurs with respect to meaning experiences. Just as there is an auditory perception of an expression's sounds plus a certain morphosyntactic organization that is utterly disconnected from also grasping a meaning in that expression, there is, too, a vision of a picture's vehicle colors and sounds plus a certain 2D organization of them that is utterly disconnected from also grasping a pictorial significance in that vehicle.

Yet, and this is my fundamental point here, things change once one grasps a *further, 3D*, organization level in the picture's vehicle. If one also sees the vehicle's 3D higher-level organizational properties, one cannot fail also to see a scene *in* it.<sup>6</sup> For then one has a seeing-in experience whose CF counts as a sufficient condition for its RF (provided that, as I said before, the seeing-in experiencer masters the relevant concepts that enter into the content of the RF of that experience).<sup>7</sup> Indeed, unlike meaning experiences, the experience

---

<sup>6</sup> Clearly enough, in order to so see the vehicle pictorially as it were, one must take a certain vantage point, as anamorphoses clearly show. But this is another issue.

<sup>7</sup> One might legitimately wonder whether this is always the case. What about *sculptorial* seeing-in, in which since the vehicle is already 3D one must not project depth on it? One might say that one three dimensionally organizes the sculpture's vehicle without seeing anything in it. Yet even in that case, the proper visual grouping affecting the 3D elements of the vehicle as a *sculpture's* vehicle in the CF of the relevant seeing-in experience

of the scene presented by the picture in the RF of a seeing-in experience is not juxtaposed to the original vision of the picture's vehicle. Instead, the original vision of the vehicle, as an (admittedly structured) 2D object among other physical objects, is preliminarily turned into a CF of a seeing-in experience; namely, into the vision of a proper *pictorial* vehicle, i.e., as something structured in a 3D organization. This CF makes that RF available.

By means of suitable examples, we can appreciate this point step by step. First of all, it is fairly clear that one may see 2D objects in a mere 2D organizational way without having any pictorial experience with respect to them. Consider the Mach figure (fig. 1). One can see this 2D object either as a diamond or as a tilted square by simply changing the organization of its elements with respect to its 2D structure. As a result, one has two perceptual experiences corresponding to the different readings of that figure along with different 2D organizations. Yet, no pictorial experience is involved here.

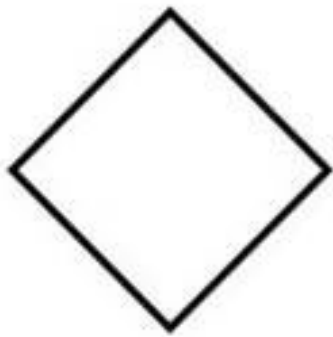


Fig. 1

Yet in addition, one can see an already organized 2D figure in a further 3D way; namely, as a *picture*, thereby no longer entertaining a vision of it in its mere 2D organization, but having a proper seeing-in experience. This second way of seeing the figure admittedly mobilizes a

---

enables the vision of that vehicle to be suitably modified, so that another 3D scene emerges in it as grasped in the RF of that seeing-in experience. This may be easily captured in Luca Patella's *Vasa physiognomica*, a 3D sculptorial instance of the Rubin's vase. In order for one to see a 3D 'facial' organization as protruding out of a receding 3D background organization, so that in a 3D vase and its surrounding space one sees two faces in profile as standing out of a background, one must suitably rearrange visually that vase and that space.

phenomenal shift with respect to the mere 2D vision of that figure in a mere 2D-based organization. But there is more than this. In this new seeing-in experience, the vision of the 2D figure itself is modified. For it now amounts to a CF of that experience grasping the figure in a new 3D organization, which lets the RF of that experience emerge, in which a 3D scene is grasped. For example, one can certainly see the Kanizsa triangle (fig. 2) in a mere 2D-based organization, by seeing a big white triangle as (partially) surrounded by three smaller triangles and three pacman-like black figures. Here, no pictorial experience of the relevant figure occurs. One has a mere 2D experience of that figure; namely, an experience of the same kind as that had with the Mach figure. Yet secondly, one may also, and more naturally, have a seeing-in experience with that figure, by taking it as a picture of a white triangular body partially occluding both another such body and three circular black bodies. In that experience, the 3D scene involving such bodies is experienced in its RF by means of a suitable 3D perceptual reorganization of the figure's elements in its CF. That CF, therefore, no longer coincides with the vision of that figure as a mere 2D organized structure.

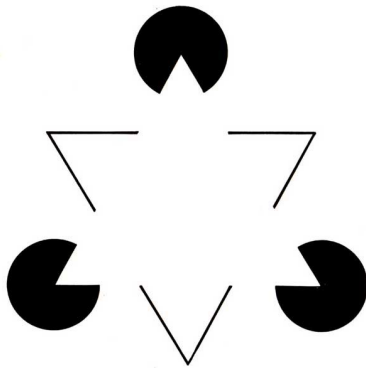


Fig. 2

By means of this example, one can understand how in its presentation of a 3D scene, the RF of a seeing-in experience occurs by virtue of suitably modifying the vision of a mere 2D figure into a vision of a silhouette suitably arranged in a 3D organization. This latter vision immediately counts as the CF of that very seeing-in experience. Thus, unlike a meaning experience, the RF of a seeing-in experience is not juxtaposed to a previous 2D perceptual experience of the figure. Instead, it emerges out of the modification of that experience, where such a modification amounts precisely to the CF of that seeing-in experience.



Now, in the case of the Kanizsa triangle one may alternate between the nonpictorial and the seeing-in experience. Either one has, admittedly less naturally, a mere nonpictorial vision of the figure, or one has, more naturally, a seeing-in experience of it, in whose CF the vision of the picture's vehicle is suitably rearranged in terms of a 3D organization. But in other cases, the seeing-in experience may occur *after* the nonpictorial experience, thereby vividly showing how the relevant picture's vehicle is then differently experienced perceptually. This happens with so-called 'aspect dawning' pictures, such as the picture of a Dalmatian (fig. 3). Unlike the Kanizsa case, the pictorial reading of the figure is less natural than its nonpictorial reading.<sup>8</sup> Only when the 2D black-and-white spots of that picture's vehicle are visually experienced in such a way that some of them are grouped together as standing *in front* of some others taken as a *background*, does one also experience in that vehicle a 3D scene with a Dalmatian dog standing in front of a background. The original nonpictorial experience of the figure is modified in terms of the CF of a seeing-in experience. This fold interpenetrates with the RF of that seeing-in experience, in which this scene is grasped. Indeed, unlike meaning experiences, the RF does not add itself to a previous lower-level experiential aspect, but it occurs by virtue of that CF, i.e., by virtue of the modification of a previous 2D vision of a picture's vehicle – taken as a mere 2D object among other physical objects – in terms of the vision of that vehicle *qua* a 3D *pictorial* vehicle.



Fig. 3

---

<sup>8</sup> On this difference between the Kanizsa case and the Dalmatian case (and the fact that cognitive penetration is necessary for the nonpictorial and the pictorial reading respectively), cf. Zeimbekis 2015.

All in all, therefore, unlike a meaning experience, a seeing-in experience involves the fact that its CF is a modification of the vision of the picture's vehicle taken in isolation. Thus, the RF of that experience does not add itself to its CF, but it is intimately connected with the visual modification that CF amounts to.

I should like to explain further this difference between a meaning experience and a seeing-in experience by stressing that, unlike a meaning experience, a seeing-in experience is a *recognitional* experience, i.e., an experience in which (like other perceptual experiences) one recognizes what its subject is. In other words, it is an experience as of an *F*, where *F* is a sortal – e.g., *being a Dalmatian* (Lopes 1996).<sup>9</sup> (Schier (1986) captured this difference by claiming that, unlike linguistic understanding, pictorial understanding is qualified by *natural generativity*. Once one understands an expression of a language, one cannot thereby understand any other expression of that language unless one knows its conventional meaning. By contrast, once one understands a picture (by virtue of recognizing its subject), one can thereby understand indefinitely many other pictures, provided that one manages to recognize their subjects.)

In fact, one cannot read off from hearing an expression, however morphosyntactically organized, what its meaning is. Lexically ambiguous expressions (e.g., “bank”) clearly show this point. Keeping in one's auditory experience its morphosyntactic organization fixed, one can experience such an expression either with a certain meaning or with another.

Yet one can read off from seeing a picture's vehicle, when properly 3D organized, what it presents. Both ‘aspect dawning’ pictures and perceptually ambiguous pictures (the duck-rabbit picture, the Rubin's vase, etc.) show this point. In ‘aspect dawning’ pictures, once one properly three dimensionally organizes the elements of the relevant picture's vehicle, one can see something in it, but not something else utterly different: e.g. a Dalmatian dog, but not a rhino. In perceptually ambiguous pictures, once one properly three dimensionally organizes the elements of the relevant picture's vehicle either one way or the other, one can respectively see in it two different somethings, but not other somethings that have nothing to do with such reorganizations. For example, in accordance with two different 3D groupings, in the duck-rabbit picture one can see either a duck (on a background) or a rabbit (on a background), but not an elephant (on a background). Likewise, in accordance with two different 3D groupings, in the Rubin's vase picture, one can see either a white vase on a

---

<sup>9</sup> Pace Wittgenstein, who once said that meaning is a physiognomy (2009<sup>4</sup>:§ 568).

black background or two black faces in profile on a white background, but not a white giraffe on a black background.

Let me clarify this difference between meaning experiences and seeing-in experiences in a further way. Although both in a meaning experience and in a seeing-in experience the higher-level aspect existentially depends on the lower-level aspect, only in the latter case does the second aspect supervene on the first (provided the experiencer masters the concepts that constitute the RF's content). Once again, one may experience different meanings in an ambiguous expression that one auditorily groups in the same way. So, there is no supervenience of the higher-level meaning aspect of a meaning experience on its lower-level aspect concerning the auditory perception of the relevant expression. Yet as perceptually ambiguous pictures show, (conceptually recognized) changes in the RF of the relevant seeing-in experience match changes in the 3D organizational grouping of the vehicle's elements seen in the CF of that experience (while the lower-level perceptual properties of the picture's vehicle may remain the same, as happens precisely in such pictures). Thus, the RF of a seeing-experience supervenes on the CF of that experience.<sup>10</sup>

So, the first difference between meaning experiences and seeing-in experiences is that only in the latter case, from sensuously experiencing its lower-level aspect one can read off its higher-level aspect. Let me now go on to consider the second difference between meaning experiences and seeing-in experiences. This difference depends on the first. In its way of being a recognitional experience, unlike a meaning experience a seeing-in experience is a genuine *perceptual* experience, though *sui generis*. This is as it must be; otherwise, if it were a mere perception-based yet basically imaginative experience, our kind of grasping e.g. of puzzle pictures as Wollheim (1987) noted, as Wollheim (1987) noted. Consider what

---

<sup>10</sup> If one distinguishes what a picture presents from its subject, once this is taken as what a picture is about (Husserl 2006, Nanay 2016, 2018, AUTHOR 1-2), one cannot read off from the picture's vehicle properly threedimensionally organized what that picture is about, even if this is constrained by what the picture presents. Indeed, a picture may present the same scene, hence its seeing-in experience still mobilizes the same RF (and also the same CF), and yet be possibly or actually about different things compatible with that scene. This makes the picture not perceptually, but merely representationally ambiguous (Wiesing 2010, AUTHOR 1). For example, by still seeing in it the very same serious man in a hieratic pose, an incompetent audience may take Piero's fresco of St. Louis de Toulouse as being about Michael Schumacher, the former F1 pilot, instead as being about St. Louis. Or a snapshot of Alan Parker's *Evita*, in which one sees an elegant lady on a certain background, may count both as a picture of Madonna, the actress playing the role of Evita Peron in that movie, and as a picture of Evita herself. But for the purposes of this paper I leave this point aside.

happens when one links dots on a paper so that all of a sudden a (depicted) turtle experientially pops up. This popping up indicates a perceptual character of that experience.

First of all, the CF of a seeing-in experience has surely a perceptual character: it is the (almost entirely: see immediately below) veridical perception of the picture's vehicle. But the RF of that experience, too, has a perceptual character. Indeed, as Levinson (1998) originally suggested, the RF of a seeing-in experience amounts to a knowingly illusory perception of the scene the picture presents: the picture's vehicle is seen, knowingly illusorily, as that scene. This perception is determined by the proper 3D grouping, in the CF of the same seeing-in experience, of the picture's vehicle elements on an illusory third dimension projected on these elements. Indeed, this grouping enables that scene to emerge, hence to be recognized. For example, in the picture of a Dalmatian, a Dalmatian dog perceptually emerges in front of a background by virtue of the fact that one groups some of its vehicle's black-and-white spots by means of a subjective contour, which lets these spots appear in front of other such spots now appearing behind. Thus, a Dalmatian dog is recognized as standing in front of a background.

Moreover, the specificity of seeing-in as a perceptual experience consists in the fact that, unlike standard perceptual experiences, that scene is not felt as present in the knowingly illusory aspect that the RF amounts to. For in the CF, the veridical fold of that experience that makes the RF knowingly illusory, *the vehicle* is instead felt as present (Matthen 2005, Dokic 2012, AUTHOR 1, Ferretti 2018). Yet seeing-in remains a perceptual experience. For, as in any perceptual experience, in the RF of a seeing-in experience the scene is still seen as present, albeit knowingly illusorily, thereby featuring the RF's content. In other terms, in the RF the picture's vehicle is seen (knowingly illusorily) as a scene being out there.

By contrast, even if it is based on the auditory perception of an expression in its morphosyntactic structure, the overall meaning experience is not a perceptual experience, not even a *sui generis* one. For since its proper meaning higher-level aspect is *not* a perceptually-driven recognitional experience, the overall experience is indeed an experience, but it is not a *perceptual* experience. At most, it is a *cognitive* experience, involving just a form of the so-called *cognitive* phenomenology (cf. Bayne-Montague 2013, Chudnoff 2015, Kriegel 2016, Montague 2016, Pitt 2014).<sup>11</sup>

---

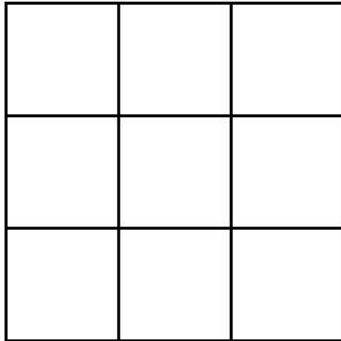
<sup>11</sup> Granted, one might wonder whether there is a *sui generis* cognitive phenomenology, or whether that phenomenology can be reduced to a kind of sensuous phenomenology (Carruthers-Veillet 2011, Prinz 2011, Tye-Wright 2011). If this were the case, one might surmise that the phenomenology of a meaning experience

Let us see more in detail why the overall meaning experience is not a perceptual experience. To be sure, in a meaning experience the proper meaning higher-level aspect is given *immediately*, just as the object of a perceptual experience. Once one has the proper linguistic mastery, one experiences that meaning automatically; one does not choose to experience it (Pinker 1994). Yet, immediacy is merely a necessary, but not sufficient condition for an experience to be perceptual (Nes 2016). In particular, this is the case if that experience is also, in its lower-level aspect, an experience of grasping, admittedly perceptually, higher-level properties of the item experienced – in this case, the morphosyntactic properties of an expression. For, unlike seeing-in experiences, the holistic attentional work that allows higher-level morphosyntactic properties to be perceptually grasped cannot be exploited in order to perceptually pick up meaning properties in the proper meaning higher-level aspect of the meaning experience. Let me expand on this.

In the CF of a seeing-in experience, first of all, attention holistically works in order to perceptually grasp the higher-level 3D organizational configuration of the relevant elements in the picture's vehicle. 'Aspect dawning' pictures show this point, but perceptually ambiguous pictures show it even better. As Jagnow (2011) points out, in the following grid one can see either an X-shaped structure in front of a background, or a cross-shaped structure in front of another background. Yet if one focused one's attention just on adjacent cells in the grid, one could not entertain this sort of Gestalt switch. Rather, one must entertain a holistic form of attention that simultaneously relates a certain cell either to certain other cells or to some other, different cells.

---

remains perceptual. Yet, there may be different meaning experiences of one and the same lexically ambiguous expression such that not only that expression remains auditorily experienced in the same way as far as its morphosyntactic identity is concerned, as we saw before, but also it involves the same mental imagery. For more on this, cf. Martina-Voltolini (2017).



Moreover, attention performs this holistic work in the CF of a seeing-in experience, thereby enabling a scene to emerge in the RF of that experience, hence making seeing-in a recognitional experience. Indeed, by virtue of that work one immediately perceives (though knowingly illusorily) in that RF what the picture presents. To come back to ‘aspect dawning’ pictures, by holistically grouping the vehicle’s black-and-white spots in the appropriate 3D way, one sees a Dalmatian dog to be out there, even if one knows that it is not there.

Now, in the lower-level perceptual aspect of a meaning experience, one holistically groups an expression’s morphosyntactical organization via attention as well. By different holistic attentional arrangements of the following famous sentence from *Mary Poppins*, in perceiving it, one parses it as having different readings:

(6) Speakin' o' names, I know a man with a wooden leg named Smith

namely, the trivial reading of a man named “Smith” and the comical one of a man whose wooden leg is so-called.

Yet, by merely relying on that attentional work, one does not manage to perceptually grasp the relevant expression’s meaning properties in the proper meaning higher-level aspect of that experience. That meaning must be provided in another way (typically, by convention). Consider again the previous meaningless (5). Once one has morphosyntactically grouped its expressions either in one or in the other way – admittedly, by relying on holistic attentional arrangements in perceiving (5) – one does not capture yet the meanings that in those

groupings (5) respectively possesses. For there are no such meanings, until (5)'s components are conventionally provided with a meaning. Thus, unlike a seeing-in experience, as regards a meaning experience the attentional work performed in its lower-level perspectual aspect on the expression concerned does not make its higher-level aspect perceptual in order to induce any recognition of its meaning. Hence, it fails to make the whole experience perceptual as well.

To be sure, one may acknowledge that the fact that the meaning experience is an experience *as of* understanding shows that it may be nonveridical (O' Callaghan 2011, but also Strawson 1994:7). Indeed, one might have that very experience and still the relevant expression might fail to have that meaning, or might have no meaning at all. (This may happen when someone competent in one language hears a sentence in a foreign language whose words, albeit only acoustically, recall words of her own language. "Paddle your own canoe" may be heard by French speakers as "Pas d'elle yeux Rhone que nous".<sup>12</sup>)

Yet by itself, this nonveridicality does not make the higher-level meaning aspect of that experience a *perceptual* illusion, hence the meaning experience is not perceptual as a whole. There are also other nonperceptual experiences that are accompanied by possibly nonveridical feelings. Take the feeling of location that constitutively occurs in an interoceptive experience. People with a limb amputated may still feel pain in that no longer existing limb. Yet since there is no such limb, the feeling is wrong. But even if it essentially accompanies the sensation of pain, that feeling does not make the sensation perceptual.

Let me take stock. I have shown that there are two important differences between meaning experiences and seeing-in experiences: 1) unlike meaning experiences, in seeing-in experiences one can read off the higher-level aspect from its lower-level aspect; 2) unlike meaning experiences, seeing-in experiences are perceptual experiences (though *sui generis*). To be sure, one might reply that, if, as we saw in the previous Section, two important similarities are not enough to let those experiences be experiences of the same kind, not even two important dissimilarities are enough in order to make them experiences of different kinds. Yet, those dissimilarities are more relevant than those similarities. For such dissimilarities make it the case that, unlike meaning experiences, seeing-in experiences are recognitional experiences of a sort that also makes them perceptual. Now, it is independently assessed in the literature that experiences differing as regards their being perceptual vs. nonperceptual differ in their *mode*, and that such a mode is enough in order

---

<sup>12</sup> I owe this example to an anonymous referee.

to make them typologically different (e.g. Searle 1983, Crane 2001, Chalmers 2004). Ditto for meaning experiences and seeing-in experiences. That is, in their being (distinctively) perceptual, seeing-in experiences are typologically different from meaning experiences.

#### 4. *The dissimilarities between meaning experiences and seeing-in experiences: 2*

Given the above dissimilarities, meaning experiences and seeing-in experiences are better construed as not the same kind of experience. But there is more than that. The deepest reason lurking behind such differences that makes such experiences typologically different is that, unlike a meaning experience, a seeing-in experience, given the relationship between its aspects, amounts to a fusion experience in a *substantial* sense: namely, a *proper* fusion experience. Let us see why.

As Stumpf (1890) underlined by talking of hearing a chord of notes or seeing a flock of birds, a proper fusion experience is an experiential whole that does not coincide with its experiential parts taken in isolation, for when such parts figure in that whole, they interact with each other. As we saw before, in a meaning experience its lower-level perceptual aspect consisting in the auditory perception of an expression along with its morphosyntactic organization remains the same as the auditory perception of that expression so individuated taken in isolation. Indeed, it still merely grasps the expression's sounds and their morphosyntactic organization. Thus, this experience is *not* a proper fusion experience.

Yet, seeing-in is an experiential whole that definitely does not coincide with of its experiential parts taken in isolation, precisely because when such parts figure in that whole, they interact with each other as its CF and RF respectively. As Wollheim (1987) said, the CF and the RF of that experience are so interconnected that, although both have a perceptual character, neither of them is identical with the corresponding perceptual experience taken in isolation: respectively, the vision of the mere picture's vehicle *per se* and the face-to-face vision of the scene the picture presents.<sup>13</sup> In the light of what we saw above, I may now say what constitutes the perceptual difference between, on the one hand, the CF and the vision of the picture's vehicle *per se* and, on the other hand, the RF and the face-to-face vision of the picture's subject. On the one hand, unlike the CF, the perception of the mere vehicle

---

<sup>13</sup> As regards the picture's vehicle, Husserl (2006) shares this idea: in picture perception but not in its perception *per se*, the vehicle is seen as a *pictorial* vehicle. Cf. Eldridge (2018).



*per se* grasps only a 2D organization, not a 3D organization, of it. By contrast, in the CF of a seeing-in experience, the latter sort of organization renders a vehicle a *pictorial* vehicle, i.e., something suitable for allowing a different scene to emerge in that experience's RF. On the other hand, unlike the RF, the face-to-face perception of the scene presented by a picture is not cognitively penetrated, at least in a strong sense. For concepts are not needed for shaping the content of a face-to-face vision of a scene. By contrast, they are needed for shaping the content of a RF.<sup>14</sup>

Incidentally, this classification of a seeing-in experience as a proper fusion experience does not undermine its being a *sui generis* perceptual experience. According to Textor (2018), in other proper fusion experiences that for Stumpf, as we just saw, were the paradigmatic cases of the genre, i.e., hearing a chord of notes or seeing a flock of birds, their folds respectively determine an overall look enabling a sensory unity (the chord, the flock) to be distinguished from its environment. This also holds of seeing-in experiences. Though complex, a seeing-in experience is just an experiential whole of a picture whose folds determine an overall look enabling that picture to be distinguished from its environment. Yet unlike paradigmatic fusion experiences, in a seeing-in experience the folds' contribution to that look is *specific*, not collective: the CF and the RF contribute to it in their own ways. This depends on the already recalled fact that, unlike hearing notes in a chord or seeing birds in a flock, there is an existential dependence of one fold (the RF) on the other (the CF). Indeed, the RF would not provide its proper contribution to the look of a picture if the CF did not also provide its own contribution.

Let me sum up. A seeing-in experience is typologically different from a meaning experience not only because of its specific perceptual way of being a recognitional experience, but also because, unlike a meaning experience, it is a proper fusion experience.<sup>15</sup>

---

<sup>14</sup> In AUTHOR 1, I provide an argument as to why this is the case. Only in the RF, but not in the face-to-face vision of the picture's subject, one needs a conceptual content. For one needs to perceptually distinguish that subject from a *different* item, i.e., the picture's vehicle, which in the CF of the relevant seeing-in experience one co-locates where, in the RF, one knowingly illusorily sees that subject.

<sup>15</sup> This paper has been originally presented in a seminar at Institut Jean Nicod, April 4 2018, Paris, and in two other conferences: *The Borders of Perception*, The Van Leer Jerusalem Institute and the Hebrew University of Jerusalem, January 16-17 2018, Jerusalem; *Themes from Aesthetics and Phenomenology*, Department of Philosophy and Education Sciences, University of Turin, June 7 2018, Turin. I thank all the participants for their stimulating remarks. I warmly thank Giulia Martina and Elisabetta Sacchi for their detailed comments on previous version of the paper, as well as two anonymous reviewers of this Review for their helpful criticisms.

## References

AUTHOR 1-2

- Bayne, T. and Montague, M. (2011). "Cognitive Phenomenology: an Introduction". In T. Bayne and M. Montague (eds.), *Cognitive Phenomenology*. Oxford: Oxford University Press, 1-34.
- Bantinaki, K. (2018). "Stylistic Deformity and Pictorial Experience". In J. Pelletier and A. Voltolini, *The Pleasure of Pictures*. London: Routledge, 125-143.
- Brown, J. (2010). "Seeing Things in Pictures". In C. Abell and K. Bantinaki (eds.), *Philosophical Perspectives on Depiction*. Oxford: Oxford University Press, 208–236.
- Briscoe, R. (2018). "Gombrich and the Duck-Rabbit". In M. Beaney, B. Harrington, and D. Shaw (eds.), *Aspect Perception after Wittgenstein: Seeing-As and Novelty*. London: Routledge, 49-88.
- Carruthers, P. and Vellet, B. (2011). "The Case Against Cognitive Phenomenology". In T. Bayne and M. Montague (eds.), *Cognitive Phenomenology*. Oxford: Oxford University Press, 35-56.
- Chalmers, D. (2004). "The Representational Character of Experience". In B. Leiter (ed.), *The Future of Philosophy*. Oxford: Oxford University Press, 153-181.
- Chudnoff, E. (2015). *Cognitive Phenomenology*. London: Routledge.
- Crane, T. (2001). *Elements of Mind*. Oxford: Oxford University Press.
- Dokic, J. (2012). "Pictures in the Flesh: Presence and Appearance in Pictorial Experience". *British Journal of Aesthetics* 52: 391–405.
- Eldridge, P. (2018). "Depicting and Seeing-in. The 'Sujet' in Husserl's Phenomenology of Images". *Phenomenology and the Cognitive Sciences* 17: 555–578.
- Ferretti, G. (2018). "Visual Feeling of Presence". *Pacific Philosophical Quarterly* 99: 112–136.
- Gombrich, E. (1960). *Art and Illusion*. London: Phaidon.
- Hagberg, G. (2016). "Leonardo's Challenge. Wittgenstein and Wollheim at the Intersection of Perception and Projection". In G. Kemp and G.M. Mras (eds.), *Wollheim, Wittgenstein, and Pictorial Representation*. London: Routledge, 117-159.
- Hopkins, R. (1998). *Picture, Image and Experience*. Cambridge: Cambridge University Press.
- Hopkins, R. (2008). "What Do We See in Film?". *The Journal of Aesthetics and Art Criticism* 66: 149–159.
- Husserl, E. (2002). *Ideas: General Introduction to Pure Phenomenology*. London: Routledge.
- Husserl, E. (2006). *Phantasy, Image Consciousness, Memory*. Dordrecht: Springer.
- Jagnow, R. (2011). "Ambiguous Figures and the Spatial Contents of Perceptual Experience: A Defense of Representationalism". *Phenomenology and the Cognitive Sciences* 10: 325–346.

- Kriegel, U. (2015). *The Varieties of Consciousness*. Oxford: Oxford University Press.
- Levinson, J. (1998). "Wollheim on Pictorial Representation", *The Journal of Aesthetics and Art Criticism* 56: 227–233.
- Lopes, D. (1996). *Understanding Pictures*. Oxford: Oxford University Press.
- Macpherson, F. (2012). "Cognitive Penetration of Color Experience: Rethinking the Issue in Light of an Indirect Mechanism". *Philosophy and Phenomenological Research* 84: 24–62.
- Macpherson, F. (2015). "Cognitive Penetration and Nonconceptual Content". In A. Raftopoulos and J. Zeimbekis (eds.), *The Cognitive Penetrability of Perception: New Philosophical Perspectives*. Oxford: Oxford University Press, 331–358.
- Martina, G. and Voltolini, A. (2017). "Perceiving Groupings, Experiencing Meanings", *Rivista di estetica* 66: 22–46.
- Matthen, M. (2005). *Seeing, Doing, and Knowing*. Oxford: Oxford University Press.
- McDowell, J. (1998). "In Defense of Modesty". In *Meaning, Knowledge, and Reality*. Cambridge MA: Harvard University Press, 87–107.
- Montague, M. (2016). *The Given*. Oxford: Oxford University Press.
- Nanay, B. (2016). *Aesthetics as Philosophy of Perception*. Oxford: Oxford University Press.
- Nanay, B. (2018). "Threefoldness". *Philosophical Studies* 175: 163–182.
- O' Callaghan, C. (2011). "Against Hearing Meanings". *The Philosophical Quarterly* 61: 783–807.
- Nes, A. (2016). "On What We Experience When We Hear People Speak". *Phenomenology and Mind* 10: 58–85.
- Pinker, S. (1994). *The Language Instinct*. Harmondsworth: Penguin.
- Pitt, D. (2004). "The Phenomenology of Cognition, or, What Is It Like to Think That P?". *Philosophy and Phenomenological Research* 69: 1–36.
- Prinz, J. (2011). "The Sensory Basis of Cognitive Phenomenology 1". In T. Bayne and M. Montague (eds.), *Cognitive Phenomenology*. Oxford: Oxford University Press, 174–196.
- Schier, F. (1986). *Deeper into Pictures*. Cambridge: Cambridge University Press.
- Searle, J.R. (1983). *Intentionality*. Cambridge: Cambridge University Press.
- Strawson, G. (1994). *Mental Reality*. Cambridge MA: The MIT Press.
- Stumpf, C. (1890). *Tonpsychologie*, vol. II. Leipzig: Hirzel.
- Textor, M. (2018). "Apprehending a Multitude as a Unity". In S. Lapointe (ed.) *Philosophy of Mind in the Nineteenth Century*. London: Routledge, 190–213.
- Tye, M. and Wright, B. (2011). "Is There a Phenomenology of Thought?". In T. Bayne and M. Montague (eds.), *Cognitive Phenomenology*. Oxford: Oxford University Press, 326–344.
- Wiesing, L. (2010). *Artificial Presence*. Stanford: Stanford University Press.
- Wittgenstein, L. (1977). *Remarks on Colour*. Oxford: Blackwell.
- Wittgenstein, L. (1991). *Remarks on the Philosophy of Psychology vol. 1*. Oxford: Blackwell.
- Wittgenstein, L. (2009<sup>4</sup>). *Philosophical Investigations*. Oxford: Blackwell.

- Wollheim, R. (1980). *Art and its Objects*. Cambridge: Cambridge University Press.
- Wollheim, R. (1987). *Painting as an Art*. Princeton: Princeton University Press.
- Wollheim, R. (1998). "On Pictorial Representation". *The Journal of Aesthetics and Art Criticism* 56: 217–226.
- Wollheim, R. (2003a). "In Defense of Seeing-in". In H. Hecht, R. Schwartz and M. Atherton (eds.), *Looking into Pictures*. Cambridge MA: The MIT Press, 3–15.
- Wollheim, R. (2003b). "What Makes Representational Painting Truly Visual?". *Proceedings of the Aristotelian Society* suppl. vol. 77: 131-147.
- Zeimbekis, J. (2015). "Seeing, Visualizing, and Believing. Pictures and Cognitive Penetration". In A. Raftopoulos and J. Zeimbekis (eds.), *The Cognitive Penetrability of Perception: New Philosophical Perspectives*. Oxford: Oxford University Press, 298-325.