Abstract

In this paper we claim, first, that there are high-level visual experiences of grouping properties, i.e., properties that an array of elements we see can have to be organised in a certain way. Second, we argue that there are auditory experiences of groupings that share certain important properties with visual experiences of groupings, thereby being perceptual and high-level as well. Third, these results enable us to understand the nature and structure of our meaning experiences. We claim that, although meaning experiences depend on the auditory experiences we discuss, they are not perceptual, for they lack one of the features that make those auditory experiences perceptual; namely, being holistically attended to.

1. Experiences of visual grouping properties as perceptions of high-level properties

Among the candidates for being high-level perceivable properties, there surely are grouping properties, i.e., the properties for a certain array of elements to be arranged in a certain order (i.e., in conformity with a certain reference frame) along a certain dimension. We here focus on visual grouping properties; namely, those properties that suitably arrange a certain array of visual elements, which are primarily qualified by low-level visual properties such as colours and shapes. We claim that such grouping properties exhibit paradigmatic features of perceivable properties, while at the same time belonging to a different level than low-level visual properties.

In the case of visual experiences, grouping properties are experientially salient when either a Gestalt switch or an ‘aspect dawning’ occurs. In the first case, one experiences a change in the organisation of the visual elements one perceives. The most basic case is the change in the organisation of an array of 2D elements along the first and the second dimension. Take the
switch that affecting a 2D figure, the ‘square-kite’ figure, when what looked like a distorted square happens to be experienced as a ‘kitish’ figure, or *vice versa*. The figure itself is thus perceived as having now a certain, now another, organisation. An analogous Gestalt switch occurs in the more complex case of a change in the organisation of an array of 2D elements not only along the first and the second dimension, but also along the third dimension, leading to a change in one’s *pictorial* experience. When seeing the Necker cube, one goes from grasping a 2D figure as having a cubish pattern with a certain protruding face and a certain receding face, leading to the experience of the picture of a cube of a certain kind, to grasping that figure as having another cubish pattern with a different protruding face and a different receding face, leading to the experience of the picture of a cube of a different kind, or *vice versa*. Likewise, in the case of the Rubin vase, in which one ends up alternating the experience of a picture of a white vase on a black background with the experience of a picture of two black faces in profile on a white background.

In the second case, an ‘aspect dawning’ occurs: suddenly, a certain array of elements comes to be experienced as organised in a certain way. This typically happens in the case of arrays of 2D elements that are also grouped in the third dimension, leading again to a pictorial experience. Consider the case in which, by suddenly grouping a set of black and white spots into a figure-ground segmentation that produces a certain ‘dalmatianish’ organisation, one ends up experiencing the picture of a Dalmatian.

As the examples of Gestalt switches clarify, grouping properties are very fine-grained properties. One and the same array of elements instantiates a certain grouping organisation (in the ‘square-kyte’ case, the ‘squarish’ organisation; in the Rubin vase case, the ‘vasish’ organisation) in all possible worlds in which it instantiates another such organisation (respectively, the ‘kytish’ organisation, the ‘facish’ organisation) (Raftopoulos 2011: 508). If one is an intentionalist and likes the idea that the phenomenal character of an experience is captured by its intentional content, one may say that such properties determine the *nonconceptual* content of an organisational experience. Thus, *qua* veridical, that experience successfully represents such properties in its nonconceptual content. Granted, in some cases at least, concepts are involved in our experience of grouping properties. If we had no concept of a Dalmatian, we would hardly see a certain figure in a ‘dalmatianish’ way. Yet concepts here play just a causal, not a constitutive, role: while they prompt that grouping experience, they are not necessary for it, thereby not figuring within its content (Tye 1995: 140).

---

Granted, this fine-grainedness of grouping properties may lead one to suspect that such properties are not objective; namely, that they are not properties of the array one faces in an organisational experience, as we have supposed all along. Rather, they would be subjective properties, i.e., properties of the (possibly different) experience(s) one entertains in facing that array (Peacocke 1983: 24-26). Yet this suspicion is ungrounded. First, in a Gestalt switch, the different experiences of grouping properties rely on the very same experience of the same low-level properties (colours, shapes) of what one faces. Conversely, if one's experience of low-level properties changes, grouping properties do not eo ipso change. Change the colours and the sizes of certain spots in the Dalmatian picture and yet you still see in it a polka-dotted silhouette on a given background.² Second, if grouping were a matter of experience, then changing the array one faces in an organisational experience would not lead to a grouping change. Yet this is not the case. If one surrounds the set of 2D marks that one experiences in a 'parallelepidish' way with a jungle of other marks, one no longer sees those marks in that way (Kanizsa 1979).

Yet even if one accepts that grouping properties are objective properties, they might be strongly mind-dependent, i.e., dependent on a subjective point of view for their individuation. For much must still be settled with the nature of the order in which the elements of an array are arranged. One may claim that the order matches an orientation in an egological space, as the appeal to egological orientation points (left-right, top-down/bottom-up, and front-back) in the description of the relevant switch seemingly shows; the ordering occurs in a viewer-centred frame (Raftopoulos 2009, 2011).

But even this claim seems ungrounded. Suppose that we drew the 'square-kite' figure on a window pane and that we looked at it both from the front and from the back of the pane. Suppose also that the different ways of seeing the figure, the 'kitish' one and the 'squirash' one, occurred by essentially grouping its lines either in an upright-to-downright direction or in a downleft-to-upright direction when seeing the figure frontally. We should then conclude that when seeing the figure from behind, new further ways of seeing the figure would emerge by grouping the figure's lines in an upright-to-downleft direction and in a downright-to-upleft direction respectively. Yet in actual fact, from both sides there are two mere ways of grouping the figure. We can thus conclude that the figure

² One might retort that grouping is a matter of having different perspectives on what one faces, thereby still mobilising subjective factors of experience. Yet first, there are many accounts of perspectives in terms of objective properties, either as mind-dependent ones (Hopkins 1998, Hyman 2006) or as situation-dependent ones (Schellenberg 2008, Fish 2009). Second, if grouping were a matter of perspective, how could different groupings be given once one's perspective of what one faces is fixed? Third, changing her perspective on a certain array does not lead one to perform another grouping. Getting closer to, or distancing oneself from, some black and white spots does not alter one's 'dalmatianish' grouping of them.
is arranged according to a frame of reference, but that this is a *cardinal* frame of reference determined by objective poles (say, a northeast-to-southwest way and a northwest-to-southeast way), instead of an egological frame of reference.\(^3\)

The defender of the subjectivity of grouping properties may appeal to the fact that, for 2D items, the relevant organisational arrangement occurs along a dimension of depth that such items do not possess. Consider the Dalmatian picture: even if one sees a ‘dalmatianish’ silhouette in front of a background, the picture is still a *flat* object, a 2D item. Now, we do not want to deny that the experiences involved in such cases, which as we saw in passing are nested within pictorial experiences, exhibit a form of projectivity.\(^4\) In order for the relevant grouping operation to obtain, depth is projected onto elements that do not possess it, since they are merely two-dimensional. Yet projectivity is not necessary for an experience of grouping. For an organisational experience of grouping occurs also as far as 3D elements are concerned, where the depth of 3D elements is not merely projected, but properly perceived, as in the 3D versions of the Necker cube and the Rubin vase. Both in the 2D case and in the genuinely 3D case, fixing the 3D dimension is required merely because the ordering in accordance with a certain direction occurs with a certain *orientation*: certain elements of the relevant array are seen as *protruding* with some other elements that are seen as *receding*. In both the 2D and the 3D cases of the Necker cube and the Rubin vase, elements that were previously seen to protrude are now seen to recede, and *vice versa*.

As this discussion on the nature of grouping properties should have already shown, such properties are *high-level* properties. There could not be such properties if there were no low-level visual properties such as colours and shapes. Indeed, the former properties depend for their instantiation on the instantiation of the latter properties. Moreover, this dependence is *generic*: it is a dependence not on certain colours and shapes in particular, but on some colours and shapes or other. However, it does not amount to a supervenience relation, as Gestalt switches clearly show. By means of such switches one and the same array of elements, which instantiates certain low-level visual properties, also instantiates different grouping properties; the grouping organisation of that array changes even if the colours and shapes it possesses remain the same (Wittgenstein 2009\(^4\): II, xi, § 247).

Acknowledging that grouping properties are high-level properties because they depend on low-level visual properties does not yet mean that they are

---

\(^3\)For other criticisms of viewer-centred frames, cf. Jagnow (2011: 336-337) and Macpherson (2006: 91, 107-108). To be sure, the fact the grouping properties depend on an *ordering* may prompt one to claim that they are at least *weakly* mind-dependent properties, i.e., properties that depend on a subjective point of view for their instantiation. On this cf. Voltolini (2015); for the distinction between strong and weak mind-dependence, cf. Newall (2011: 67).

\(^4\)Spinicci (2012: 98-100) speaks of *apparent* depth.
high-level perceivable properties. Granted, both the Gestalt switches and the ‘aspect dawning’ cases show that there is an experiential difference, respectively, between passing from grasping a certain organisation to grasping another one, and when detecting such an organisation in an array that seemed not to have it. Yet if experiences, qua conscious mental states, are not exhausted by perceptions or anyway by mental states endowed with sensory features, as we will see in Sec. 3, one may wonder whether such a phenomenal change is also a genuine perceptual change. Could not that change be a change just in cognitive phenomenology,5 that is, in a phenomenology irreducible to sensory phenomenology (or, if it is so reducible, is to be typically traced back to mental imagery)?

Granted, this point shows that the so-called method of phenomenal contrast (Siegel 2011) fails to ground the idea that grouping properties are grasped in perceptual experiences, thereby being perceivable properties. E.g. in the ‘aspect dawning’ case a phenomenal contrast indisputably obtains: there is a change in phenomenal character between a ‘before’ situation, in which no relevant grouping property is grasped, and an ‘after’ situation, in which that property is grasped. But what can guarantee that that change involves a perceptual change, rather than a change in a phenomenally relevant cognitive feature of the experience?

However, we can independently claim that grouping properties are perceivable properties. First, our experiences of grouping properties share some important features with experiences of paradigmatically perceptual properties such as low-level visual properties. Consider again Gestalt switches. Grouping properties are responsible for many features that qualify such switches, namely: exclusivity (the alternate experiences in the switch are not given simultaneously), inevitability (one way of seeing the object will eventually replace another), and randomness (the duration of one alternation in that switch is not a function of previous durations) (Block 2014: 567). Now, many of such features are also typical of perceptual experiences of low-level properties like the shapes, colours, or sizes of visible objects: we do not choose e.g. to see the colours of the objects we see, nor do we attend to their sizes because we attended to their shapes. Second, attention also plays a decisive role, by being another necessary condition of the perceptual character of grouping properties. For in the experience of such properties, attention does not pre-perceptually work as a local spotlight enabling the experiencer to perceive certain areas of the scene she faces.6 Rather, it works holistically, by perceptually rearranging in an utterly different way the elements of the array one faces.7 Likewise in a Gestalt switch involving mere 2D different configurations, in order for one to see e.g. a grid of nine squares either as ordered according to a cross-organisation as opposed to an X-organisation, one cannot merely focus

5 Cf. e.g. Kriegel (2011: 157-158).
6 Cf. e.g. Pylyshyn (2003), Rafepoulos (2009).
on one of the squares: no switch will occur. Rather, one must see that square as properly connected to other squares of the grid (Jagnow 2011: 342).

We may then conclude that visual grouping properties are high-level yet genuinely perceptual properties. Thus, their experiences are genuinely perceptual. In the proceeding of this paper, we will argue that there are grouping properties that are still perceptual yet depend for their instantiation on the instantiation of low-level properties that we perceive in another sensory modality: namely, auditory grouping properties. Our experiences of these properties indeed exhibit important similarities with experiences of visual grouping properties, thereby being high-level and perceptual as well.

2. Experiences of auditory grouping properties as perceptions of high-level properties

In this section, we argue that grouping properties are not restricted to the visual domain and focus on auditory grouping properties. We first provide some evidence, both phenomenological and empirical, for the existence of these properties. Plausible examples range from organisational properties of simple sound streams to musical properties and linguistically-relevant properties of speech sounds. We then argue that our experiences of these properties are genuinely perceptual, but irreducible to experiences of more basic audible properties. We will conclude by addressing some issues concerning the nature of auditory grouping properties.

What can we experience auditorily? It is widely agreed that we can hear sounds or auditory objects, and properties such as pitch, loudness, duration, and timbre. However, there is reason to think that what is perceptually given to us in auditory experience comprises more than these basic audible properties. Grouping properties analogous to the visual grouping properties we talked about in Sec. 1 are a promising candidate. Evidence for the existence of auditory grouping properties comes from two main kinds of phenomena: complex auditory experiences, such as experiences of music or of speech, and cases of perceptual multistability. All these phenomena involve experiences whose phenomenology or qualitative character cannot be fully explained in terms of our awareness of the pitch, loudness, duration, and timbre of individual sounds.

Experiences of music involve an awareness of rhythms, melodies and harmonic progressions, which is irreducible to unorganized sounds and their basic properties. Hearing e.g. a melody is not just hearing a sequence of sounds with certain pitches (Von Ehrenfels 1988: 83-85). A melody retains its identity across transpositions, i.e., when it is played starting from a different note. Upon transposition, every pitch in the sequence will be different, but as long as the melodic intervals between the sounds are preserved, the melody will sound the same. The best explanation of this phenomenology is that we experience the
melodic relations between the individual sounds. This, we argue, is a matter of experiencing auditory grouping properties of the sounds constituting the melody. While we can only experience musical properties such as melodic relations if we experience certain auditory low-level properties, the former are irreducible to, and do not even supervene on, the latter. Analogous remarks can be made for harmonic intervals, i.e., relations between the pitches of simultaneous sounds, rhythms and metrical relations.

Experiences of speech also support the appeal to auditory grouping properties. Hearing sounds as speech involves an experience of segmented and organised strings of sounds that makes linguistically relevant units such as words and sentences salient to us. Without this organisation, speech sounds would be heard as a continuous and uniform stream, much as it happens when listening to someone speaking in a completely unfamiliar language. The phenomenal difference between hearing speech in a language one does not know and hearing speech in a language one is familiar with – or between hearing the same speech sounds in a language one is not familiar with and hearing them after becoming acquainted with it – is best explained in terms of the ability to experience phonemes, syllables, words, and possibly syntactically relevant properties of whole sentences.

Plausibly enough, at least some linguistically relevant properties are auditory grouping properties. When we experience syllables or words, that is, we experience certain grouping properties of the speech sounds we hear, which allow us to parse the string of sounds into units that are significant in our language. While we cannot experience these grouping properties unless we experience the low-level properties of the speech sounds, the former cannot be reduced to the latter, nor do they supervene on them. Experiencing e.g. a certain syllable is compatible with experiencing a wide range of low-level properties, as suggested by the various ways the same syllable sounds when pronounced by people of different age, bodily size, or in different emotional states. Moreover, as we will illustrate shortly, the same combination of phonemes is compatible with different syllables; since phonemes are themselves irreducible to low-level

As suggested by a referee for this journal, the phenomenology of hearing a melody can be explained in terms of the experience of a complex auditory object and the individual sounds as part of this object. This characterization is not in conflict with our appeal to grouping properties: experiencing the individual sounds as grouped in a certain way is what allows one to experience a higher-order object. Acknowledging the experience of a higher-order object, though, does not allow us to get rid of grouping properties. As we will see, a higher-order object such as a melody can itself be experienced according to more than one grouping, i.e., one can experience, at different times, multiple grouping properties of this object.

For arguments to the effect that this phenomenal difference is best explained by supposing that phonemes are auditorily experienced, cf. O’Callaghan (2010, 2011). For us, linguistically relevant properties other than phonemes can also be auditorily experienced. We remain neutral on whether phoneme experience is a matter of experiencing grouping properties, and focus on syllables and words instead.
auditory properties (O’Callaghan 2010: 322-323), the grouping properties responsible for experiences of syllables are a fortiori irreducible to such low-level properties. Properties of speech sounds that account for our experience of words and perhaps even phrases, clauses and sentences are other potential examples of auditory grouping properties in the linguistic domain.

Cases of auditory multistability provide further support for auditory grouping properties. Just like experiences of visually ambiguous figures, experiences of ambiguous auditory objects offer phenomenal contrasts that can be best explained by supposing that we experience different grouping properties before and after the Gestalt switch. Here is one compelling case. Certain strings of speech sounds can, when constantly repeated over an interval of time, constitute multistable auditory objects. In the so-called verbal transformation effect, the same stimulus can give rise to a «life life life» percept, which alternates over time with a «fly fly fly» percept (Warren and Gregory 1958). The low-level auditory properties remain constant, but our experience changes with respect to what it is like to have it. In order to account for this phenomenal change, we need to suppose that some property of the auditory object other than a low-level auditory property is being experienced after the Gestalt switch that was not perceived before. The best explanation, we claim, is that certain auditory grouping properties are experienced; thus, the sounds and their basic auditory properties are organised into different sequences of syllables.

We have argued that experiences of music, speech, and ambiguous auditory objects support the thesis that we can experience auditory grouping properties, i.e. that our awareness of these properties makes a distinctive phenomenal difference. However, as we saw in Sec. 1, one may legitimately doubt that the experiential contrast between, say, hearing sounds and hearing them as speech, or hearing the «life life life» organisation as opposed to the «fly fly fly» organisation is a perceptual contrast. In fact, we have not yet established that the difference in phenomenology pertains to perceptual – here, auditory – as opposed to cognitive phenomenology. We will now try to show that experiences of music and speech as well as experiences of multistable auditory objects share some key features with paradigmatically perceptual experiences and that, therefore, auditory grouping properties are perceptual properties.

We have independent reasons for thinking that experiences of certain musical and linguistically relevant properties are perceptual. First, it is not necessary to be a trained musician or an expert in music theory in order to be able to experience melodies, harmonic progressions and rhythms (DeBellis 1991). We know that infants are sensitive to octave equivalence and sensory consonance and dissonance; moreover, they can recognise melodies and remember them more


Cf. e.g. Demany and Armand (1984); Schellenberg and Trainor (1996); Trainor and Heinmiller (1998).
accurately than individual pitches. These abilities are best explained if we suppose that infants, children and non-expert adults are sensitive to melodic intervals. Since by hypothesis these subjects do not possess music theoretic knowledge or recognitional capacities for melodic intervals, but only the perceptual capacity to discriminate them, it is plausible that their sensitivity is perceptual. Analogous empirical findings also speak in favour of the perceptibility of rhythms, metres (Hannon and Johnson 2005, Ladinig et al. 2009), and harmonic relations – all of which are irreducible to low-level properties of sounds.

Second, experiencing linguistically relevant grouping properties does not require possession of the concepts for those properties or previously acquired knowledge of the language at stake. Not only do infants exhibit a preference for speech sounds as opposed to nonspeech and are sensitive to differences between different languages, but they also start developing the ability to segment continuous speech sounds into audible units with linguistic significance very early on. Nothing in the low-level properties of speech sounds can determine such units, hence where word boundaries, pauses and phrase boundaries are.

If infants can have experiences of segmented and organised speech sounds, this suggests that they perceive high-level properties as well. Perceiving these grouping properties allows them to perceive speech sounds as organised into syllables and then words, and plausibly into clauses and phrases.

Some theorists doubt our claim that experiences of the grouping properties involved in music and speech sounds do not require recognitional capacities and do not depend on previously acquired knowledge of music or language. While they acknowledge that such experiences are perceptual, they insist that, at a minimum, familiarity with a certain musical environment and style, or some degree of knowledge of a certain language is necessary. Arguments for this view are often based on empirical evidence suggesting that infants have different discrimination capacities for musical or linguistically relevant properties than adults. Prior to consistent exposure to a certain language or musical system, infants would be sensitive to fine-grained audible differences, which are not specific to that language.

12 Cf. e.g. Trainor and Trehub (1992); Plantinga and Trainor (2005, 2009).

13 In experiments testing the ability to detect tonality and therefore to harmonic relations, subjects with no musical training show performances that are very similar to those of musical experts. Cf. e.g. Trainor and Trehub (1994), and 4-5 year olds already show sensitivity to tonality; cf. e.g. Trainor and Trehub (1994); Corrigal and Trainor (2009); Krumhansl and Keil (1982); Sloboda (1985); Cuddy and Badertscher (1987).

14 Cf. e.g. Vouloumanos and Werker (2007); Bosch and, Sebastián-Gallés (1997).


16 Cf. e.g. Bartel (2007), Scruton (1999). O’Callaghan (2011) focuses on phoneme perception, but the form of his argument might apply to any language-specific audible property.
or musical culture and could be explained in terms of perception of low-level auditory properties. Once infants are familiar with a certain language or musical system, by contrast, language and musical system-specific properties become salient to them, resulting in an adult-like discrimination behavior. This evidence may plausibly indicate that knowing a language or the rules of a musical system is necessary in order to perceive the grouping properties specific to the sounds of that language or musical system. One’s experience of those properties would thus depend on non-perceptual cognitive capacities and memory.

In response to this form of argument, we want to highlight that the empirical evidence is compatible with a different reading. Rather than involving a loss of one’s discrimination capacities, being exposed to a language or musical culture makes a difference to which properties one pays attention to and which ones one disregards for the purposes of understanding and communication; moreover, it affects the ease, speed, and accuracy of one’s discrimination and recognition. Thus, speech and music perception are cognitively penetrable only in a weak sense. Granted, the phenomenology of auditory experiences of speech and music can be altered by the content of the non-perceptual cognitive states and by the concepts of the subject. However, the cognitive influence is merely causal, rather than constitutive. There are no conclusive reasons for thinking that perception of auditory grouping properties of speech sounds and music depends on concept possession, recognitional capacities, background knowledge, or even familiarity with a language or musical system. If this is true, then experiences of auditory grouping properties share an important feature with experiences of low-level audible properties such as pitch, loudness, duration, and timbre.

There is also a more general motivation for the claim that experiences of auditory grouping properties are perceptual. Auditory capacities underpinning our everyday perception are sufficient for experiencing at least some grouping properties. While our auditory system is constantly reached by many unorganised stimuli, our conscious experience usually presents us with stable, determinate

17 Some empirical research suggests that learning a language brings about a change in which similarities and differences are more salient, rather than a change in one’s perceptual discrimination capacities. These capacities, in fact, remain available. For example, adults can, though with some effort, still make very fine-grained distinctions that are not relevant in their native language; cf. Werker (1994, 1995), Kuhl et al. (2006). An analogous interpretation is also available for data on the different discriminatory behavior of infants and adults in the musical domain; cf. Hannon and Trehub (2005), Lynch et al. (1990), Trehub et al. (1999).

18 If the influence were constitutive, the subject could not have had the experience with the phenomenal character and the content that the experience in fact has if she did not entertain the specified cognitive state or if she lacked the specified concepts; cf. Siegel (2011: 205-206).

19 In addition to the capacities involved in basic auditory streaming here discussed, perceptual capacities specific to, say, speech perception and musical perception may also be required for experiences of certain grouping properties.
properties and discrete auditory objects. How can our auditory system achieve this striking result? Plausibly, a form of organisation occurs already at the most basic level of auditory experience. As research on auditory streaming has shown, perceptual organisation is guided by principles similar to those first invoked by Gestalt psychologists for the visual modality; plausibly, its function is to inform the subject about the worldly objects and properties that can be regarded as the sources of sounds. Harmonicity, onset synchrony, continuous frequency modulations are cues for ‘fusing’ together frequency components that are likely to originate from a single source, so that the pitch of a single sound is heard. Temporal proximity, pitch proximity, timbre, and good continuation also are ecologically valid criteria for grouping different sounds together into segregated patterns called streams (Bregman 1990). In an ordinary auditory scene, a car passing, a dog barking, and a human voice are segregated into autonomous streams, although such organisation is not determined by the audible stimuli alone. Even though knowledge and familiarity may influence the way the scene is experienced, they are not necessary: there is no need for the subject to conceptualise the sounds as belonging to a certain recognised source in order to hear them as a stream. Moreover, the Gestalt principles guiding visual organisation incorporate ecologically useful assumptions that are hard-wired in the perceptual system, innate, subpersonal and automatic, and already apply to the experiences of infants. Since, plausibly, the same holds of Gestalt principles in the auditory modality, we can conclude that auditory streaming is a low-level, stimulus-driven, purely perceptual process.

The principles of auditory streaming can also explain auditory multistable experiences. The so-called ‘auditory streaming paradigm’ itself generates, under certain conditions, a perceptually ambiguous auditory object. The paradigm is an «ABA_ABA_…» sequence where A and B are isochronous sounds differing in pitch and «_» denotes a silence of the duration of one of the sounds. If the pitches of the two sounds in the pattern are relatively similar, the sequence is heard as a fused melodic line with a galloping rhythm. When the difference in pitch is higher or the sequence is played at a fast tempo, so that the distance in time between the tones decreases, the sequence is perceived as two segregated streams: an «A_A_…» percept and a «_B_B_…» percept, one at twice the repetition rate of the other (Heise and Miller 1951; Bregman and Campbell 1971; Van Noorden 1975).

Here it may seem that the above auditory difference depends on a change in the pitch and duration of the sounds in the sequence. Yet interestingly enough, for certain values of repetition rate and pitch, the «ABA_» scheme generates an

---

20 Cf. e.g. Wertheimer (1938), Bregman (1990).
21 Cf. e.g. Pylyshyn (2003).
22 Cf. e.g. Quinn and Bhatt (2015).
ambiguous auditory object. Let us fix the repetition rate and pitch within this range of values. Now, when hearing this object under prolonged repetition, a multistable condition is reached and subjects experience reversals between the fused and the segregated percept. These Gestalt switches are spontaneous, involve changes in auditory phenomenal character, and have dynamics of alternation similar to those observed for visually ambiguous figures (Denham and Winkler 2006; Pressnitzer and Hupé 2006). Stimulus-driven, perceptual processes guided by low-level principles of streaming can explain the groupings occurring in the phenomenally different experiences we can have when hearing the streaming paradigm. At the same time, we need to appeal to high-level grouping properties in order to account for the phenomenal character of these experiences. Since, by hypothesis, the repetition rate of the sequence, hence the duration of the sounds, as well as the pitch of the sounds are constant, our experience of these low-level audible properties cannot explain the Gestalt switches. We have therefore found a paradigmatically perceptual experience of auditory grouping properties.

This purely perceptual account can arguably be extended to more complex auditory multistable experiences. Even our very simple «ABA_» sequence supports a further level of multistability analogous to the visual multistability of Rubin’s Vase picture. Upon hearing the two segregated streams, the As line and the Bs line can be perceptually organised according to a figure-ground relation, so that one stream can be heard as a foreground melody while the other is heard as an accompaniment. As many have noticed, the contribution of perceptual attention to our experience of the multistable object well explains the relative phenomenological salience of the foreground melody. However, as in the visual modality, the role of attention here is not to focus on one stream of sounds, allowing us to perceive more of its low-level audible properties or making our experience of them more determinate. Rather, attention performs a holistic organisation of the elements we perceive.

At the first level, when our ambiguous auditory object is experienced either as a fused stream or as two segregated streams, attention organises the low-level properties of the sounds in the sequence so that we experience two different groupings. Attention plays an analogous role at the second level of multistability, when the segregated streams are alternatively heard as foreground melody


24 The grouping activity involved in streaming seems to account for the perception of a wide range of musical properties, including those involved in polyphonic effects, which can give rise to multistable experiences (cf. e.g. Dowling 1973; Bregman and McAdams 1979; Deutsch 1999; Huron 2001; Pressnitzer et al. 2011).

25 Cf. e.g. Bregman and Campbell 1971; Hupé and Pressnitzer 2012; Moore and Gockel 2012.
and accompaniment. Here the streams – which in turn result from more basic groupings – are grouped according to a certain figure-ground relation; thus, a holistic organisation of the complex auditory object constituted by the two streams is experienced. Again, the outcome of this attentional activity is not a more determinate or more focused experience of the same low-level properties of the auditory object: attention does not have a merely post-perceptual role, since it allows us to perceptually experience properties of the auditory object that we could not perceive before, i.e., certain grouping properties.

That attention plays a key role in bringing the experience of grouping properties about does not undermine the claim that such experiences are perceptual. Attention can modulate the influence of nonperceptual cognitive states such as intentions, beliefs, and desires on experiences of grouping properties. However, one’s perception of such properties is largely independent of one’s nonperceptual states. In the multistable case, for example, one never has complete voluntary control on the temporal features of the alternation of the different grouping experiences, nor can one bring oneself to experience a different grouping when one can only notice one of them – the switch will happen, if it will, spontaneously (Van Noorden 1975; Pressnitzer and Hupé 2006). What is more, one cannot auditorily experience the sounds one hears as organised according to whatever pattern: which low-level properties of an auditory object one grasps constrains which perceptual groupings are possible. For example, when one experiences an ambiguous string of speech sounds as organised according to a «life life» grouping and then according to a «fly fly» grouping, this is because the phonemes one hears are compatible with more than one segmentation into syllables. Given those phonemes, one could not have an auditory experience of, say, a «bly bly» syllable grouping – or, better, if one did, one would be merely imagining.

While these remarks speak in favour of the perceptibility of auditory grouping properties, do they tell us something about their nature? We argued in Sec. 1 that visual grouping properties are objective and only weakly mind-dependent, but one may doubt that the same holds in the auditory domain. We know that experiences of auditory grouping properties depend on experiences of low-level auditory properties in the sense that one cannot have the first kind of experience without having the second one. Moreover, we know that the low-level properties of an object put some constraints on what range of grouping properties it can have. Indeed, changing the low-level properties of an object can change its high-level grouping properties – as demonstrated by the effects of tempo and pitch variation on the multistability of the streaming paradigm discussed above.26 However, as many other experiences of ambiguous auditory objects also

---

26 As we saw before, for some values of pitch and tempo the auditory streaming paradigm is not multistable, i.e., it cannot be heard as having two different perceptual organisations.
show, changing the low-level properties is not necessary in order to generate a difference in the high-level properties. We must thus conclude that auditory grouping properties, just like their visual counterparts, do not even supervene on low-level auditory properties.

Nonetheless, we do not need to construe auditory grouping properties as subjective, i.e., as intrinsic properties of the experiences we have of them. Phenomenological considerations on experiences of music and speech support this point. These experiences seem transparent to one: when one introspects one's experience, one is seemingly presented with objects and their apparent properties, rather than with properties belonging to one's mental states. Intuitions, however, are less clear when it comes to multistable experiences. On the one hand, when one undergoes a Gestalt switch, it may seem to one as if one is discovering something about the object, rather than about one's experience or perceptual capacities. On the other hand, it does not seem to one as if the object itself is changing, as it may be when the pitch or volume of an auditory object changes. Rather, one may think, how one experiences the same object changes; nothing in one's experience indicates that that change amounts to one's discovery of some property of the object one did not perceive before. If this is true, then intuitions about the transparency of one's experience should not carry much weight.

In support of the objectivity of auditory grouping properties we can then appeal to two facts, concerning the intersubjectivity of these properties, and the factors determining them. First, there is no indication that experiences of grouping properties vary widely across different subjects. For example, in all the experiments cited in this section subjects reported the very same groupings: while they experienced the different groupings at different times and variation rates, no one auditorily experienced an idiosyncratic grouping of the same low-level properties. Second, the low-level properties of an object – which we consider to be paradigmatically objective – constrain what auditory grouping properties the object instantiates, hence what can be heard when perceiving it. By contrast, facts about the psychology of the individual perceiver do not play such a role: they may constrain whether one is able to experience a certain property – whether low- or high-level – that is there to be experienced, but not which ways of grouping the perceived object are available. These considerations suggest not only that auditory grouping properties are not properties of subjects or their auditory experiences, but also that these properties do not need to be individuated in terms of our experiences or other subjective factors.

We hope to have shown that there are grouping properties that depend for their instantiation on low-level auditory properties, and range from the very basic groupings involved in auditory streaming to musically relevant and linguistically relevant properties. Moreover, we can have experiences, as well as multistable experiences, of such properties that are both phenomenologically distinctive and genuinely perceptual. Our next step is to apply these results to
the philosophical debate on meaning experiences. We will provide an account of such experiences by investigating their relation to experiences of auditory grouping properties as well as their difference from these experiences.

3. Meaning experiences as sui generis doubly high-level experiences

Armed with the previous reflections, let us consider meaning experiences. As we will see, although they are experiences of high-level properties depending on the auditory experiences of high-level grouping properties, they are not perceptual experiences.

Meaning experiences are cases of grasping certain semantic properties, thereby enabling one to understand certain linguistic expressions. These experiences are distinctive: experiencing an expression as having a certain meaning is not the same as experiencing an expression as having another meaning.\(^{27}\) Granted, the experience of the meaning of an expression depends on a previous auditory experience of that expression:\(^{28}\) one could not experience the meaning of the expression if one did not hear it. Theoretically speaking, there may be other ways for one to understand that expression: one may e.g. be given an explanation of its meaning. However, these ways of understanding an expression do not count as experiencing its meaning, for they do not depend on hearing it.

As with experiences of grouping properties, moreover, this dependence is generic: one may have the same meaning experience even when hearing different, typically synonymous, expressions. Another similarity with experiences of grouping properties is that this dependence is intertwined with an irreducibility of meaning experiences to the auditory experiences of the relevant expressions. For they do not supervene on the latter experiences. Phenomenal contrast cases indeed reveal this irreducibility. In such cases, one entertains different meaning experiences and still has the same auditory experience of the very same expression. Experiences involving lexical ambiguity paradigmatically show this. Suppose that one hears the sentence «Dionysius is Greek» twice, yet one experiences it the first time as meaning the nationality of Dionysius the Elder, ruler of Syracuse, and the second time as meaning the (same) nationality of Dionysius the Younger, son of the preceding. In this case, one entertains different meaning experiences even if one still has the very same auditory experience affecting the

\(^{27}\) Cf. Pitt (2004). Meaning experiences are also labeled experiences of as-of understanding: Strawson (1994). This labeling allows one to see that such an experience may be illusory (O’Callaghan 2011: 791-792): in it, one seemingly grasps a meaning for an expression even if the expression has no such meaning.

\(^{28}\) If written linguistic expressions are at stake, meaning experiences will depend on previous visual experiences of such expressions. In this paper, however, we will focus on the case of speech.
very same sentence. One-sentence puns are more vivid examples. In order for one to get the joke in «Santa’s helpers are subordinate Clauses», one must hear the sentence in the different meanings related to the different meanings of the word «clauses».

Hence, an experience of the meaning of an expression does not lie at the same phenomenal level as an auditory experience of that expression. Still, one may object that a meaning experience must be phenomenally reduced to an experience involving just an auditory experience. This experience may be a perception or even an imagination of low-level audible properties – typically, the sounds involved in hearing a certain linguistic expression. But it may even be an experience of high-level auditory grouping properties: notably, an experience that, by capturing either the phonetic units that morphologically determine single expressions or even the syntactical organisation of such expressions, ex hypothesi involves perceptual grouping operations of the kind we saw in Sec. 2.

Yet the previous “Dionysius” example rules out this objection. First, the auditory low-level perception or imagery that is involved in both meaning experiences is the same: in such experiences, one not only hears, but also imagines hearing exactly the same sounds. Second, the auditory high-level experience is also the same. For in the two meaning experiences, one ex hypothesi auditorily groups in the very same way the only sentence one hears not only with its constituent expressions, but also with its syntax.

In this respect, experiences involving lexical ambiguity differ from experiences involving structural ambiguity. For only in the latter cases there is a one-to-one correspondence between meaning experiences and the ex hypothesi high-level auditory experiences of the morphosyntactic units of the expressions involved. Indeed, in such cases one first auditorily groups the same sentence in different morphosyntactic ways, and only then one experiences different meanings with the different groupings one performed. To see this, note that one may syntactically parse in a different way even a lexically meaningless sentence. The meaningless «The slithy toves gyred the Jabberwock in the wabe» may be

---

29 For similar examples, cf. Siewert (1998); Horgan and Tienson (2002); O’Callaghan (2011).
30 We refer to this kind of experience as a high-level auditory experience of the morphosyntactic units of an expression. For the hypothesis that meaning experiences are auditory experiences in this sense, cf. Carruthers and Veillet (2011: 52); Prinz (2011); Tye and Wright (2011).
31 The low-level visual imagery involved is also the same: in the two meaning experiences, one entertains the mere image of a (past) man.
33 A similar case occurs with a mere difference in a morphological grouping. Suppose that, instead of ending up with auditorily grouping the word «life» after having iteratively heard the word «fly» (cf. Sec. 2), one ends up auditorily grouping the meaningless word «libe» after having iteratively heard the meaningless word «bly».
parsed as ‘saying’ either that in their being in the wabe, the slithy toves gyred the Jabberwock, or that the slithy toves gyred the Jabberwock in its being in the wabe. But only once the meaningless lexical items involved in the parsed sentence are given certain semantic interpretations can one entertain different meaning experiences. (Theoretically speaking, if one of such items turned out to be lexically ambiguous, one might entertain up to four meaning experiences, respectively doubling the two different syntactic parsings one already performed with the involved sentence.)

This is why we relied on lexical ambiguity to illustrate the thesis that meaning experiences both depend on, and are irreducible to, high-level experiences of auditory properties; namely, groupings concerning expressions at the morphological level and possibly also at the syntactic level. Since with syntactic ambiguity, as we just saw, there is a one-to-one correspondence between meaning experiences and experiences of the morphosyntactic units of expressions, one may be tempted to deny that experiences of the former kind involve a further phenomenal level over and above experiences of the latter kind. Yet this temptation can be resisted by stressing, as we just did, that with lexical ambiguity there is no such correspondence.

However, one might still be skeptical as to whether the contrast occurring in grasping a lexical ambiguity is a phenomenal contrast. Granted, in that case one grasps different meanings. Yet why must such a grasping be experientially loaded? This doubt can be dispelled by noticing that a phenomenal contrast occurs even when there is a one-to-one correspondence between a meaning experience and an experience of the morphosyntactic units of the expressions involved. For in this case, one can see that the former state is a meaning experience irreducible to an experience of the second kind. Consider a situation in which one experiences a meaningless sentence that one merely morphosyntactically groups in a certain way and another situation in which one experiences that sentence as also having a certain meaning. Note that it may be even the case that one is first independently given the meaning of the sentence one just morphosyntactically groups in a certain way, and only later on one also experiences that sentence as having that meaning. Suppose that someone experienced the famous sentence «Pirots carulise elatically» à la Carnap, i.e., as a syntactically well-formed yet meaningless sentence. Suppose further that our subject learnt that such a sentence actually has a meaning and that meaning were explained to her. This still would not be enough to bring about a phenomenal change in

34 This case is analogous to the phenomenal contrast case that Pitt (2004: 28-29) and Reiland (2015: 485-487) respectively point out. All such cases involve no difference in high-level perceptual groupings between the first and the second experience.

35 Nes (2016: 72-77) proposes a similar sophistication of the phenomenal contrast case. Yet he unnecessarily resorts to an example involving a foreign language that one subject only grasps morphosyntactically while another subject fully understands it.
her experience. But the subject would undergo that change once she experienced the sentence as having that meaning – say, that air pilots (what «pirots» ends up meaning) twirl around (what «to carulise» ends up meaning) elegantly (what «elaticy» ends up meaning).\textsuperscript{36}

So far, it has turned out that meaning experiences are experiences of high-level properties – semantic properties – that depend on perceptual, still high-level grouping properties of linguistic expressions but are irreducible to them. But what sort of experiences are they? The fact that, as we just saw, one can appeal to phenomenal contrast cases in order to show that there are meaning experiences may lead some people to claim that they are perceptual experiences, just as both experiences of low-level auditory properties and experiences of high-level auditory grouping properties are.\textsuperscript{37} Yet, as we already saw in Secs 1-2, the fact that the phenomenal contrast method is applicable to a certain class of experiences does not yet mean that such experiences are perceptual. The phenomenal difference that method appeals to between morphosyntactically grouping a sentence and experiencing it as having a certain meaning may be a difference between a perceptual and a cognitive form of phenomenology.\textsuperscript{38} Granted, there is a good reason to say that meaning experiences are perceptual: such experiences exhibit the same kind of automaticity, or more generally immediacy, as other experiences we classify as perceptual, notably experiences of grouping properties (as we saw in Secs 1-2).\textsuperscript{39} Yet there is also a good reason to question the claim that mean-

\textsuperscript{36} Once this phenomenal contrast case is so conceived, it is not vulnerable to the objection raised against Strawson’s original example (1994: 5-6) of an English monoglot who does not understand a French sentence. According to this objection, the phenomenal difference between the meaning experience a Frenchman has and the experience the English monoglot has upon hearing the French sentence is a just a difference between a low-level and a high-level auditory experience. Unlike the English monoglot, the Frenchman does not hear mere sounds, but morphosyntactically groups the sentential elements in a perceptually relevant way (cf. Carruthers and Veillet 2011: 52; O’Callaghan 2011: 799). Whether this objection is effective against Strawson’s case or not, it does not apply to our case. In our case, the phenomenal contrast lies between an auditory, admittedly high-level, experience in which one groups a sentence just up to the syntactic level and another experience in which one also understands it. Likewise with a case of satiation in which repeatedly hearing a word results in a meaningless experience of it. For satiation also occurs in the case of the iteration of a meaningless word: the iterated word is no longer perceptually grouped as such, but is heard as a mere sound. Unlike the former case, the latter case only mobilises a phenomenal contrast between a high-level auditory experience of morphological groupings and a low-level auditory experience of sounds.

\textsuperscript{37} Cf. e.g. Bayne (2009), Siegel (2011). For a recent defense of this claim, cf. Nes (2016).

\textsuperscript{38} As O’Callaghan (2011: 806) also suggests. See also Reiland (2015: 491), who however remains neutral between a liberal interpretation of cognitive phenomenology, on which it is a sui generis form of phenomenology, and a conservative interpretation of it, on which it is reduced to some other perceptual form of phenomenology. We will now choose the first option.

\textsuperscript{39} Nes (2016: 77-81) nicely summarises the reasons why meaning experiences should be classified as perceptual. Among such reasons, Nes mentions the sensory adaptation satiation
ing experiences are perceptual. Immediacy is not a sufficient condition for an experience to be perceptual, although it is a necessary condition (Nes 2016: 83).

To begin with, since meaning experiences depend on, but are irreducible to, experiences of auditory grouping properties that in their turn depend on, but are irreducible to, experiences of low-level auditory properties, one may doubt that meaning experiences are perceptual. Even if one allows for high-level experiences of grouping properties to be perceptual, as we did in Secs 1-2, how can meaning experiences, \textit{qua doubly} high-level experiences, be also perceptual?

In itself, this doubt is not very problematic. For being doubly high-level does not in principle prevent an experience from being perceptual. Indeed, being merely high-level is perfectly compatible with being a perceptual experience. In Sec. 2 we saw that high-level experiences of phonemes, syllables, words and possibly even syntagms are all perceptual, although they instantiate different levels of complexity.

The main problem is another. As we saw in Secs 1-2, the other necessary condition for an experience of high-level properties to be perceptual is that they are attended to \textit{holistically}. Indeed, the groupings that induce distinct morphosyntactic orderings in hearing the very same expression mobilise this holistic way for attention to work.\footnote{As Carruthers and Vellet (2011: 52) seem to acknowledge.} Yet no such work occurs in order for an expression to be experienced in a certain semantic way, a fortiori in different such ways. One may hear «Pirots carulise elatically» as having a certain meaning only once one ascribes meaning to the expressions constituting it. Unlike its morphosyntactic grouping, no holistic attentional work affecting its sentential elements occurs in experiencing that sentence as meaningful. A fortiori, no such work occurs in differently experiencing the meaningless sentence «The slithy toves gyred the Jabberwock in the wabe» as ultimately having different meanings. Granted, the fact that that sentence may be differently parsed is a matter of holistically attending it differently. Yet as we saw before, this perceptually relevant attentional work operates at the morphosyntactic, not at the semantic level. Finally, in differently hearing a lexically ambiguous sentence, such as the «Dyonisius»-sentence above, no different such work occurs.

It is time for us to reach two main conclusions. First, not only there is a (doubly) high-level meaning experience having a certain phenomenal character, as the phenomenal contrast method shows, but also that character is not perceptual, but cognitive. Thus, meaning experiences are not perceptual; they involve cognitive phenomenology.

Granted, one may wonder whether a meaning experience can really be an \textit{experience}, if it is not perceptual. What sort of difference is there between enter-
taining that experience and hearing the relevant expression plus being given its meaning (Nes 2016: 74-75)? Well, unlike the latter case, the former case involves a dependence of the meaning experience on the auditory experience of the relevant expression. This already supplies the mental state of grasping the meaning of the expression with an experiential character. Moreover, this dependence does not require that that state is perceptual. A realisation of a mathematical proof is experiential because it depends on perceptual experiences of suitably arranged concrete elements, but it is not perceptual (Chudnoff 2015: 116).

Second, since a meaning experience depends on, but it is irreducible to, a high-level auditory experience of grouping properties, in particular of a morphosyntactical type, its cognitive phenomenal character is irreducibly nonperceptual. Thus, we opt for a liberalist approach to the cognitive phenomenology of meaning experiences.41

References

Bartel, C.

Bayne, T.

Block, N.

Bosch, L., Sebastián-Gallés, N.

Bregman, A.

Bregman, A., Campbell, J.

Bregman, A., McAdams, S.

41 Although this paper has been jointly conceived and discussed, Giulia Martina is specifically responsible for Sec 2, Alberto Voltolini is specifically responsible for Secs 1 and 3.
Carruthers, P., Vellet, B.

Chudnoff, E.

Corrigal, K., Trainor, L.

Cuddy, L., Badertscher, B.

DeBellis, M.

Demany, L., Armand, F.

Denham S.L., Winkler I.

Deutsch, D.

Dowling, J.

Fish, W.

Heise, G.A., Miller, G.A.

Hannon, E., Johnson, S.

Hannon, E., Trehub, S.

Hopkins, R.

Horgan, T., Tienson, J.

Hupé, J.M., Pressnitzer, D.

Huron, D.
Hyman, J.  

Jagnow, R.  

Jusczyk, P.W., Friederici, A.D., Wessels, J., Svenkerud, V.Y., Jusczyk, A.M.  

Jusczyk, P.W., Aslin, R.N.  

Kanizsa, G.  

Kriegel, U.  

Krumhansl, C., Keil, F.  

Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., Iverson, P.  

Ladinig, O., Honing, H., Haden, G., Winkler, I.  

Lynch, M.P., Eilers, R.E., Oller, D., Urbano, R.C.  

MacPherson, F.  

Moore, B., Gockel, H.  

Nanay, B.  
– 2011, *Do we see apples as edible?*, “Pacific Philosophical Quarterly”, 92: 305-322.

Nes, A.  

Newall, M.  

O’Callaghan, C.  


Orlandi, N.  

Panneton, R., Newman, R.  
Peacocke, C.

Pitt, D.

Plantinga, J., Trainor, L.

Pressnitzer, D., Hupé, J.M.

Pressnitzer, D., Sued, C., Shamma, S.

Prinz, J.

Pylyshyn, Z.

Quinn, P., Bhatt, R.

Raftopoulos, A.

Reiland, I.

Schellenberg, E., Trainor, L.

Schellenberg, S.

Scruton, R.

Siegel, S.

Siewert, C.

Sloboda, J.

Spinicci, P.
Stokes, D.

Strawson, G.

Trainor, L., Heinmiller, B.

Trainor, L., Trehub, S.

Trehub, S., Schellenberg, E.G., Kamenetsky, S.B.

Tye, M.

Tye, M., Wright, B.

Van Noorden, L.

Voltolini, A.

Vouloumanos, A., Werker, J.F.

Von Ehrenfels, C.

Warren, R., Gregory, R.

Werker, J.F.
Wertheimer, M.

Wittgenstein, L.