Amodal completion in vision and audition. A relationship between perception and mental imagery

Elvira Di Bona

Abstract

Amodal completion is nearly always discussed in terms of vision. In this paper I focus on hearing. I argue that just as we seem to amodally complete the hidden parts of a visual object through visual mental imagery, we also seem to amodally complete the masked parts of a sound through auditory mental imagery. But there remains an important difference between the two modalities.

Keywords: amodal completion, mental imagery, auditory perception, visual perception, sound

1. *Introduction*

The problem of amodal completion is how we represent an object as a whole when we do not receive any sensory stimulation from its hidden or occluded parts (Nanay 2010: 240, 2023: 56). Nearly all discussions of amodal completion address this issue in terms of vision. For instance, suppose you are looking at a cat behind a picket fence, but the cat's tail is hidden by one of the pickets. The problem is how you can still perceive the cat as a whole, despite not receiving sensory input from the hidden part.

The same problem arises in hearing. For example, imagine hearing the wail of a siren, but part of the sound is masked by a sudden burst of louder noise. The problem is how we perceive the continuous sound of the siren as a whole, even though the masked portions provide no sensory input.

In this paper, I will argue that we complete the masked sounds through mental imagery, but that the way this works in audition differs from how it works in vision. First, I will explain the view that amodal completion in vision occurs through mental imagery by focusing on the relationship between what is perceived and what is sensorily imagined, which is central to an imagery-based account of amodal completion. Then, I will turn to audition and show that, on the one hand, there are sufficient similarities to the vision case to argue that amodal completion occurs through mental imagery; and on the other hand, there is an important difference between vision and audition.

Analysis Vol. XX | Number XX | XX 2025 | 1–8 doi: https://doi.org/10.1093/analys/anaf006

My approach will be phenomenological rather than empirical. I follow Smith (2018) in defining the relevant sense of 'phenomenology' as the study of the 'appearances of things'. In this sense, Amy Kind's discussion of visual imagination is an example of a phenomenological argument, because in order to claim that imaginative presence explains amodal completion, she focuses on how certain things appear similarly in both perceptual and imagistic experiences (2018). While I share Kind's phenomenological approach, I will differentiate my discussion from hers, as she focuses on the individual mental abilities of perception and mental imagery rather than the specific relationship between them.

In contrast, Bence Nanay's arguments about amodal completion in vision exemplify an empirical approach, as they are based on a psychological and neuroscientific understanding of mental imagery (2023: 12–14).

While Kind does not discuss audition, Nanay extends his view to sensory modalities beyond vision, assuming that his conclusions about vision apply to other senses, although he does not fully commit to the details. In contrast, I will focus specifically on the case of audition. An imagery-based account of auditory amodal completion has not yet been explored in the literature on hearing.

2. *Visual perception and visual mental imagery*

We can understand amodal completion in the visual case as a consequence of some familiar observations about our experiences of visual imagery. When we imagine a specific object, it seems as if we visualize it 'in the mind's eye'. And, although this is not the simplest form of visualization, we might also be able to place the imagined object within our egocentric space (Briscoe 2011: 153, Martin 2002: 410), such as when we decide where to position an imagined painting on a wall. We simply visualize the painting as being in a specific location outside our mind, that is, on the wall. Regarding egocentric visualization, Briscoe notes:

> Human beings and perhaps other animals also have the ability to 'augment' reality by superimposing mental imagery on the visually perceived scene. Thus, as Rick Grush observes, one might use a mental image to decide where in egocentric space a vase should be placed in order best to obscure a picture on one's desk (2004: 390). Similarly, when deciding how to arrange the furniture in a new home, one might imaginatively project an armchair into an empty corner or a painting onto an unadorned wall. (2011: 153)

The latter, for instance, involves locating the painting on a specific portion of the wall where we intend to place it. Significantly, both objects – the imagined painting and the perceived wall – must be seen in the foreground. If either is relegated to the background and, thus, appears blurry, we cannot properly

relate the two objects to accurately determine where the painting should be hung. Seeing both objects in the foreground is a necessary condition for establishing their relationship, as high resolution and detailed perception are required to successfully complete the task.[1]

It is no coincidence that the same happens in ordinary perception, when no visual imagery is involved, as when someone moves a painting on a wall and asks where the painting should be hung: to properly perform the task, we need to see both the part of the wall where the painting should be hung and the frame of the painting in the foreground. Should one of them be in the background, we could not perform the task with the right precision. The parallel occurs because, at least in these ordinary cases, we experience the relationship between an imagined object and a real one to be sufficiently similar to the relationship between two real objects. That is, in using visual imagery to decide where to place a painting on a wall or furniture in a room, we take for granted that there is enough similarity between visual perception and visual imagery in the content they represent that the task can be completed successfully.[2] The important point with respect to amodal completion concerns the role of foreground and background, which is therefore worth clarifying further.

Firstly, the fact that what is in the foreground is seen in high resolution, with a higher level of detail and clarity, while what is in the background is seen in low resolution – less sharp and poorer in details if compared to high-resolution counterparts – depends on the role of consciousness. This means that while a subject can be fully aware of the attended object when it is in the foreground, the same subject is not completely aware of it when that object is in the background. Or, better, when an object is in the background, it is seen blurrily. Then, when attention brings it to the foreground, it is immediately into focus. We consciously attend to the object which is into focus and we can consciously experience what is unattended in a characteristically peripheral way (Watzl 2011: 156). Indeed, the phenomenology of

1 For my general argument what matters is just that the painting and wall are related (and see to be so), and not the specific relation they are in. However, it is possible to be more precise. Depending on the context, the nature of the visual objects, and their spatial characteristics, the relationship might involve interaction, causation, part-whole relationship, spatial proximity, and so on, as the example shows. Notice, however, that 'foregrounded versus backgrounded' does not count as a genuine relationship between the painting and the wall, since it is purely subjective.

2 Of course actual imagistic capacities vary widely, ranging from e.g. expert painters who might be able to imagine an apple's shape and colour as indistinguishable from a real apple, to aphantasics who cannot produce mental imagery at all. For present purposes I rely only on the observation, made by many philosophers of perception (e.g. Briscoe 2011, Martin 2002, Nanay 2010) and ordinary people, that at least some of us deploy imagery in performing such familiar tasks.

attending to something 'consists in having an experience that is structured into foreground and background' (Watzl 2011: 155; see also Husserl 2004).

Then, it is by virtue of the perceptual Gestalt principle of figure-ground segregation that one can determine what belongs in the background and what belongs in the foreground. This principle helps the eye differentiate an object from its surroundings, and, naturally, while both the figure and the ground are perceived, the former is seen in the foreground, with greater detail, while the latter is perceived in the background, with much less detail.

There are two ways to apply the principle of figure-ground segregation: the first depends on the interests of the subject, while the second can be applied independently of the subject's motivations. An example of the first application is the scenario already discussed, where we intentionally hang either an imagined painting or a real one on a wall. Here, we must find the right place for the painting, so we intentionally direct our attention in such a way that the most relevant things for our task (namely, the painting and the specific portion of the wall) appear in the foreground. An example of the second way to apply the principle of figure-ground segregation – which does not rely on the subject's motivations – is the Rubin's vase illusion. In this scenario, we cannot help but see the scene as either the vase in the foreground with the two faces in the background or, conversely, the two faces in the foreground with the vase in the background. We apply the principle unconsciously. In amodal completion, the Gestalt principle is applied in this second way, unconsciously.

In amodal completion, we represent an object as a whole by imagining the relationship between what we actually see and what we merely imagine. Specifically, thanks to our ability to superimpose mental imagery onto our visually perceived scene, we automatically imagine that the hidden part of an object – such as the imaginary tail of a cat – is in relationship with the visible part of that object, the cat's body, to form a complete visual representation of a cat. Both parts of the object are perceived in the foreground, therefore perceived with greater detail and precision, as amodal completion based on mental imagery is a form of figure-ground segregation that occurs independently of the subject's motivations. Regardless of what the subject intends to do, they cannot help but see the whole cat in the foreground, with the rest of the perceived scene – the surroundings of the cat – relegated to the background.

3. *Auditory perception and auditory mental imagery*

Let us now turn to the case of hearing. The question is whether, when we experience the auditory continuity of sounds, we represent the masked parts of sounds by virtue of auditory imagery, analogously to the case of vision.

The auditory *illusion of continuity* is the phenomenon of auditory amodal completion in which we are able to hear a specific sound as continuous even

though part of it has been masked by a louder sound (Bregman 1990: 28). The problem is how one perceives the continuous sound of, say, a siren wail interrupted by the ringing of a bell as a whole since, in its masked parts, this sound sends no sensory information to the experiencer. To solve the problem of auditory amodal completion, we can consider the experience of auditory mental imagery more generally. It seems that we can imagine a sound as located 'in the mind's ear', as when we imagine the slamming of a door. Similarly to vision, it also seems possible to imagine a sound as located in egocentric space. Moreover, just as we superimpose mental imagery onto the visually perceived scene when we imagine a visual object in our egocentric space, we superimpose mental imagery onto the auditory scene – the soundscape (Schaffer 1994) we are immersed in – when we imagine a sound in our egocentric space. For example, we can imagine the ringing of a doorbell coming from the direction of the entrance of our house.

Furthermore, it also seems possible to imagine the relation between an imaginary sound and a real sound, though with a difference compared to the case of the visual imagery. In vision, I can see a real wall and, at the same time, imagine a painting to be located somewhere on the wall. That is, I see the relation between two co-present objects (or parts of them) as being in the foreground; as we have seen, this is necessary in order to perform the task of correctly locating the painting on the wall. In the case of audition, instead, when I hear a siren wail and simultaneously imagine the ringing of a bell in relation to the siren wail, it seems that these two sounds (the real one and the imagined one) cannot both occupy the foreground, making it difficult to ascertain the relations between them.[3] For instance, I may want to compare them to determine which sound lasts longer. To do this, both sounds would need to occupy the foreground, allowing me to evaluate their duration accurately. While we can be surrounded by many simultaneous sounds, we can perceive only one sound in high resolution – that is, the one that occupies the foreground. Even though many sounds can populate the soundscape, the foreground of the latter is filled up completely with a single sound, while all other sounds, if present, recede into the background. Unlike the case of two co-present visual objects, I cannot ascertain the relationship between two simultaneous sounds to each other because I cannot perceive both in high resolution. To grasp the relationship between two sounds, I must hear them consecutively – that is, one after the other.

3 As with visual objects, in audition it is also possible to identify specific relationships between two sounds, such as the one in the example. Depending on the context, the nature of the sounds, and their temporal occurrences, they might be related through causation, part-whole relation, succession, identity relation, spatial vicinity and so on. However, what matters to my argument is, again, just that two sounds are related (and heard to be so), rather than the specific relationship between them.

Consider the example of a one-year-old baby learning to walk: the sound of a sudden tumble might be followed by the sound of loud crying. You can determine whether the first causes the second (or whether the first lasts longer than the second) because, being consecutive, both are perceived in the foreground. This allows them to be experienced in high resolution and rich in detail. The listener's interest determines whether both sounds are brought to the foreground, as attention directs the focus of perception. Now consider hearing the sound of a tumble and then a pause; most of us, at least those of us who have spent time with toddlers, will be prompted to *imagine* the sound of crying even if, for some reason, it does not actually follow. The perceived sound and the imagined one can, like the two real sounds, both be foregrounded because they are consecutive.

As with the vision case, the parallel occurs because we experience the relationship between an imagined sound and a real one to be sufficiently similar to the relationship between two sounds. That is, in imagining the cry following the sound of the tumble, we take for granted that auditory perception and auditory imagery are similar in the content they represent – or at least, similar enough for the task at hand (i.e. predicting what we will hear next).[4]

Also analogously with vision, in audition, what appears in the background and what appears in the foreground depends on the application of the Gestalt principle of figure-ground segregation. This application can either be influenced by the subject's motivations or occur independently of them. An example of the former is the previously discussed scenario in which we intentionally evaluate the difference in duration between an imagined sound and a real one. An example of the latter – where the application is not based on the subject's motivations – is automatically imagining the sound of crying after hearing the sound of the toddler tumbling. Auditory amodal completion via mental imagery is another such example.

As we have seen, when we listen to more than one sound – or when we listen to one sound and imagine another – the foreground of the soundscape is exhausted by a single sound (whether real or imagined) when it occupies the foreground. To perceive the relationship between two real sounds, or between a real sound and an imaginary one, these sounds must be experienced as consecutive rather than simultaneous, so that their relationship can be brought into foreground. In auditory amodal completion, sounds are perceived as occurring consecutively. This happens because we can superimpose mental imagery onto the auditory scene, enabling us to imagine the masked part of a sound as being related to the audible part, thereby forming a complete sound. Both parts of the sound are brought into the foreground, allowing

4 As with vision, auditory imagistic capacities vary widely, ranging from e.g. sound engineers who might be able to imagine sounds indistinguishably from real ones, to those who suffer from anauralia (Hinwar & Lambert 2021), who cannot produce auditory imagery at all.

them to be perceived as related and contributing to a unified auditory experience. Moreover, since amodal completion relies on mental imagery – which is a form of figure-ground segregation occurring independently of the subject's intentions – listeners cannot help but perceive the sound as amodally completed, regardless of their intentions.

Consider the example of a siren wail masked by a burst of noise. Despite the masking sound, we inevitably perceive the siren as continuous because the burst is replaced in our perception by a mental imagined fragment of the siren that connects with the audible portion. Amodal auditory completion is, therefore, a form of temporal completion in which mental imagery plays a crucial role, allowing us to complete the masked part of the sound.[5] This temporality distinguishes it from the simultaneity of visual amodal completion.

4. *Conclusion*

While in vision we can imagine an object that simultaneously appears to be in a relation with a real object – since both can be seen in the foreground – in audition, we can imagine a sound in relation to a real sound only if they occur consecutively. That is because auditory relationships can only be heard in the foreground when sounds are sequential. In vision, amodal completion is based on a simultaneous relationship between the visible parts of an object and its imagined parts, whereas in audition, amodal completion relies on the temporal relationship between the audible parts of sounds and their imagined parts.

Phenomenological considerations thus support the general claim that amodal completion in both vision and audition occur through mental imagery, but suggest greater caution in assuming that more specific conclusions that apply to one sense modality must also apply to others.[6]

Funding

This work was supported by the FIS Starting Grant (FIS00000243).

University of Turin
Italy
elvira.dibona@unito.it

5 Nanay and Young's work (2022) focuses on olfactory amodal completion, where they make a passing note that auditory amodal completion is temporal. Similarly, in a paper on perceptual objects, O'Callaghan (2008) briefly suggests that the temporal profiles of auditory stimuli are crucial to amodally completing an occluded auditory stream.

6 I am particularly grateful to Davide Bordini, Edoardo Bronzolo, Daniele Cassaghi, Stefano Ercolino, Mariaenrica Giannuzzi, Mauro Lenti, Niccolò Nanni, Donald Oxtoby, Alberto Voltolini, an Editor of this journal and the anonymous referees for their insightful comments.

References

Bregman, A.S. 1990. *Auditory scene analysis: the perceptual organization of sound*. Cambridge, MA: The MIT Press.

Briscoe, R. 2011. Mental imagery and the varieties of amodal perception. *Pacific Philosophical Quarterly* 92: 153–73.

Grush, R. 2004. The emulation theory of representation: motor control, imagery and perception. *Behavioral and Brain Sciences* 27: 377–96.

Hinwar, R.P. and A.J. Lambert. 2021. Anauralia: the silent mind and its association with aphantasia. *Frontiers in Psychology* 12: 744213.

Husserl, E. 2004. *Wahrnehmung und Aufmerksamkeit: Texte aus dem Nachlass (1893-1912)*, eds. T. Vongehr and R. Giuliani. Dordrecht: Springer.

Kind, A. 2018. Imaginative presence. In *Perceptual Presence*, eds. F. Dorsch, M. Nide-Rumelin, and F. Macpherson, 165–80. New York: Oxford University Press.

Martin, M. 2002. The transparency of experience. *Mind & Language* 17: 376–425.

Nanay, B. 2010. Perception and imagination: amodal perception as mental imagery. *Philosophical Studies* 150: 239–54.

Nanay, B. 2023. *Mental Imagery*. New York: Oxford University Press.

O'Callaghan, C. 2008. Object perception: vision and audition. *Philosophy Compass* 3: 803–29.

Schaffer, M. 1994. *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester: Destiny Books.

Smith, D.W. 2018. Phenomenology. In *The Stanford Encyclopedia of Philosophy* (Summer 2018 edn), ed. E.N. Zalta, <https://plato.stanford.edu/archives/sum2018/entries/phenomenology/>.

Young, B. and B. Nanay. 2022. Olfactory amodal completion. *Pacific Philosophical Quarterly* 103: 372–88.

Watzl, S. 2011. Attention as structuring the stream of consciousness. In *Attention: Philosophical and Psychological Essays*, eds. C. Mole, D. Smithies and W. Wu, 145–173. New York: Oxford University Press.