CHAPTER 15

ALGORITHMIC SYNESTHESIA

NOAM SAGIV, ROGER T. DEAN,
AND FREYA BAILES

We often think about our senses as separate and independent of each other. However, there are at least four problems with this naïve view. First, our experience of the world around us is a rich multisensory one. Under normal conditions, our senses are constantly flooded with information. Even when we are sitting quietly at home and reading this book, we not only see the text, but also hear the pages flipping, experience the smell of the book, and feel the smooth texture of the pages. Likewise, while drinking from a cup of tea, we would notice its smell and color in addition to the liquid's taste, temperature, and texture or viscosity. Although we can try to focus our attention on some things and ignore others, it would be safe to say that our sensory experience is rarely limited to one modality even if we try very hard.[1]

Second, our sensations do seem to come together somehow into a unified experience. The sights and sounds of most events we witness do not appear to be two independent streams of information; in fact, they seem inseparable. A third problem is that the senses do not merely converge somewhere; they also influence each other. As discussed here, visual inputs can influence auditory perception and vice versa.

The fourth reason to doubt the independence of the senses is that, for some individuals, stimulation in one sensory modality may actually give rise to perceptual experience in more than one modality. This remarkable form of perception is labeled *synesthesia*, and we distinguish here, before drawing some parallels between the two, between the involuntary psychological phenomenon and synesthesia in art involving intentional intermedia experimentation.

Bill Viola (1995, p. 152) stated: "Western science has decided it is desirable to isolate the senses in order to study them, but much of my work has been aimed at putting it back together." Scientists no longer limit themselves to studying the senses in isolation, and hundreds of recent studies have now examined cross-modal interactions and multisensory perception. However, it would be fair to say that artists were ahead of scientists on this front. Some notable examples include composers Richard Wagner and Alexander Scriabin and painters James McNeill Whistler and Wassily Kandinsky.

Wagner (1849) advocated a concept of a unified art work (*gesamtkunstwerk*, encompassing music, theater, and visual arts), which he was able to realize supervising closely the set design and writing librettos for his own operas. Scriabin was particularly preoccupied by the relationship of music and color (Galayev 1998), including a light organ in his orchestral work "Prometheus: Poem of Fire." He died before completing his "Mysterium," which was meant to include dancers and scents also.

Whistler's search for parallels between music and painting is evident in the musical titles of some of his paintings (e.g., *Nocturne* or *Symphony*; see Lochnan 2004). Likewise, Kandinsky's work was inspired by music (Betancourt 2007) (and he is in fact thought to have been a synesthete; Ione and Tyler 2003).

No doubt, technology has made it easier to create multimedia today (e.g., the simple visualization one encounters using a media player), but the central question is not how to implement it but what to implement. We soon discuss different approaches to real-time algorithmic synesthesia, in particular sharing features between simultaneously produced sound and image. We begin with the "genuine" synesthetic experience naturally occurring in a minority of individuals.

****Synesthesia* is usually defined as a condition in which stimulation in one sensory modality also gives rise to a perceptual experience in other modalities (Sagiv 2005). For example, for some synesthetes, music may evoke vivid visual imagery (e. g., colors, shapes, and textures; Steen 2001). Others may experience different sensory combinations, such as colored smell, taste, or pain (Day 2005); experience taste while listening to spoken words (Ward and Simner 2003); or have tactile experiences induced by smell (Cytowic 2002), to name a few examples. Although synesthesia covers a whole range of cross-sensory phenomena, it is often induced by lexical and conceptual inducers (e.g., letters and numbers; Simner 2007). What is common to these is that the synesthetic imagery is induced reliably, automatically, and involuntarily, and that they involve an actual perceptual experience rather than a mere association or metaphoric description (e.g., Baron-Cohen et al. 1987, Martino and Marks 2001). Although different synesthetes generally disagree on particular correspondence (e.g., B-flat may be pink for one synesthete but green for another), the mapping is consistent over long periods of time for each synesthete. This consistency is one of the hallmarks of synesthesia and often is utilized to verify the genuineness of the subjective reports of synesthetes (Hubbard and Ramachandran 2005, Rich et al. 2005).[2] Recently, psychophysical measures such as reaction times and stimulus detection rates have been used to establish consistency (e.g., Sagiv et al. 2006). Many of the experimental studies were based on congruency paradigms by which synesthetes were required to respond to stimuli colored either congruently or incongruently with their synesthetic experience (e.g., Ward and Sagiv 2007). Faster responses expected in the congruent condition were taken as corroborating evidence for synesthetes' subjective reports, although admittedly such simple behavioral measures alone may not suffice to verify the rich phenomenological descriptions provided by synesthetes.

Much of the recent literature on synesthesia has focused on the developmental form of the condition, that is, healthy individuals who have had these experiences for as long as they can remember. Nevertheless, it is possible to acquire synesthesia either temporarily (with hallucinogenic drugs or during other altered states of consciousness) or more permanently following sense organ, nerve, or brain injury (for more detailed overviews of cognitive and neural mechanisms, see, e.g., Hubbard and Ramachandran 2005, Robertson and Sagiv 2005, Sagiv and Ward 2006).

Recent studies suggested that the condition may be more common than previously thought—about 4% of the general population may experience one form or another (Simner et al. 2006). The most common inducers appear to be ordinal sequences such as letters or numbers or the days of the week. The most common resultant synesthetic experience is a color experience. Curiously enough, these forms of synesthesia do not fall within the boundaries of the common narrow definition as the inducer and concurrent percept may both belong to the same modality (e.g., a visually presented letter or word evokes a color experience). Furthermore, in some cases, thinking about the inducer may suffice to elicit the synesthetic experience (e.g., the concept of Monday evoking the color red). Nevertheless, there is a general agreement among researchers that these forms should be considered types of synesthesia in their own right (while the inclusion of other synesthesia-like variants is still debatable; Sagiv 2005).

Synesthesia was once considered a rare abnormality at best. Scientists would often dismiss odd subjective reports made by synesthetes, but attitudes are changing. We now aim to explain synesthesia rather than dismiss it. Furthermore, synesthesia can be utilized as a test case for theories of perception (surely, a good theory must be able to explain not only how a system usually works but also how it deviates from "normal" function). The remainder of our discussion of the psychological phenomenon of synesthetic perception focuses predominantly on auditory-visual interactions. Let us now introduce algorithmic synesthesia. We hope to convince you that the psychological and artistic explorations of synesthetic perception are complementary. ****Algorithmic synesthesia* is a term we introduced (Dean et al. 2006) to describe multimedia works in which sound or image share either computational process or data source. For example, a single algorithmic process, whether running freely or under the control of a performer, may generate both sonic and visual streams, ab initio or by transformation of preexistent sound and image. Alternatively, a common data set or data stream might be the source for different algorithmic processes that generate, respectively, sound and image. The core idea is that such relationships might encourage an audiovisual perceptual or cognitive interaction with some interesting features in common with synesthesia itself. There is no claim that algorithmic synesthesia necessarily involves all or even any of the specifics of synesthetic perception just summarized. However, in common with synesthetic perception, the association between audition and vision is more than metaphorical, involving shared material or algorithmic processing (see Seitz 2005 for discussion of synesthesia in relation to metaphor). It is worth noting that the sharing of underlying structure also binds a musical score (visual) to its sonic realizations (albeit with considerable imprecision), while the piano roll or its contemporary equivalent, the MIDI (musical instrument digital interface) sequence file, also binds a preconceived graphic form to a more precise sonic realization. However, unlike the situation of algorithmic synesthesia, audiences are generally not exposed to both streams of data but rather only to the sound.

Relatively few creative artists have overtly attempted algorithmic synesthesia as yet, but it has become more common (at least implicitly) because of its increasing accessibility through laptop performance. There are as yet no empirical studies of its impacts as far as we can determine (indeed, at the time of writing, Google Scholar still only referenced our work in response to a search for the term), and it has been argued more broadly that "a central concern of any new workable theory of electronically mediated meaning must be to understand the implications of multimodality" (Nelson 2006, p. 56).

1. THE COGNITIVE BASIS OF AUDITORY-VISUAL CROSS-MODAL INTERACTIONS

The significance of examining cross-modal interaction from both creative and scientific angles becomes apparent when bearing in mind that we can appreciate art, communicate about it and through it, owing to the fact that art can effectively engage perceptual and cognitive brain mechanisms common to us all (Zeki 1999). Artists are somewhat like scientists in the laboratory in that they try to probe these mechanisms and discover what drives them. Therefore, before proceeding to explore further creative aspects of algorithmic synesthesia, we review some cognitive perspectives on auditory-visual interaction.

A. Cross-Modal Convergence

As noted in the beginning of the chapter, even in the most ordinary situation, our experience of the world around us is inherently multimodal. We use information acquired by different sense organs to learn about objects and events around us. These objects and events produce familiar and sometimes less-familiar sensory signals. Our brains in turn try to make sense of these signals, identify sources, and more broadly, find meaning. This is by no means an easy task. Sensory information varies widely with context. For example, under different lighting conditions objects may reflect different wavelengths; retinal size and shape vary with distance and viewing angle; salience varies with the background and masking by other objects. Similarly, we are able to recognize invariant sound sources and patterns despite ever-changing source distance, speed, direction of motion, acoustics of the environment, and a whole range of other factors influencing the sound spectrum. Speech recognition provides a striking demonstration of our ability to extract information despite changes in identity of the speaker, accent, pitch, speed, presence of noise, and spectral distortions introduced by devices (e.g., mobile phones or poor-quality speakers).

Under normal conditions, auditory and visual information converge to facilitate perception. For example, both faces and voices can provide clues to recognizing people. When one source of information is of poor quality, the combination may be essential. Furthermore, lip reading enhances speech perception (in individuals with normal hearing, particularly in noisy environments).

Binding together visual and auditory information (e.g., in the sense of attributing them to the same source) is but one of several binding problems that our brains must solve to combine the various features we perceive into objects and make sense of a deluge of sensory data (e.g., Treisman 1996). Attentional mechanisms are thought to be essential for this process.

B. Cross-Modal Interactions

It is clear that input in one sensory modality does not merely enhance or complement perception in other modalities; it can also alter what is being perceived. For example, listening to one syllable (ba) repeatedly while seeing a person uttering a different syllable (ga) results in hearing da rather ba (McGurk and MacDonald 1976). Visual information can also override auditory information in localization tasks. The ventriloquist effect demonstrates this well: when watching a movie, you perceive speech as coming from the actor's mouth rather than the speakers to the right and left of the screen.

While vision often overrides auditory information in case of discrepancies, it is more difficult to find examples in which audition overrides vision. Shams and colleagues showed that the number of flashes of light one perceives may correspond to the number of beeps you hear during the same time window rather than the actual number of flashes displayed (Shams et al. 2000, 2002), and in recent preliminary work we have shown that the superimposition of rhythmic patterns, by means of pitch differences or to a lesser extent spatialization differences, can enhance such illusions (Wilkie et al. forthcoming). Interestingly, the ventriloquist effect can be reversed when the image becomes very blurry. Under such conditions, localizing by sound may indeed seem like an optimal choice (Alais and Burr 2004).

Such interactions suggest that multimodal information processing involves more than convergence of otherwise independent streams of information. Indeed, at the neural level, we find that activity in brain areas thought to be unimodal can be modulated by information coming from other senses (for a review, see Macaluso and Driver 2005). For example, Calvert et al. (1997) found that silent lip reading activates the auditory cortex, while Sadato et al. (Sadato et al. 1996) found that Braille reading in blind individuals is associated with primary visual cortex activation.

C. Matching and Mismatching Information

How does one determine that there is a match between auditory and visual signals? What do they have in common, and how can we compare these radically different inputs? We are looking for covariance or coupling of the signals along any two dimensions. Common onset and offset times, or generally speaking, synchronized dynamics, may imply a common source object or causal relationship. Synchrony is indeed a very powerful cue to causality or unity (even when it is improbable given physical location). The rubber hand illusion demonstrates this in the visuotactile domain (Botvinick and Cohen 1998). This illusion arises when the same tactile stimulus is simultaneously applied to the hand when it is hidden from view and to a visible rubber hand placed in an anatomically plausible position in front of participants. Under these conditions, they are likely to feel that the rubber hand is their own (i.e., visuotactile synchrony influences perceived hand position and limb ownership).

In the auditory-visual domain, the ventriloquist effect represents another illusion of unity mediated by synchrony. While the brain may tolerate a short lag,[3] the illusion breaks down at larger ones (evident in occasional broadcasts in which such a lag is present for some reason or another). Interestingly, apparent synchrony depends not only on the timing of the auditory and visual events but also on the distance of the source. For example, Alais and Carlile (2005) showed that we can tolerate and in fact expect larger lags for more distant sources, compensating for the later arrival of sound waves relative to light waves. This would relate well to the fact that film sound designers routinely use such timing differences in the belief that they accentuate the impression of reality (sounds of doors opening, etc.).

It is important to remember that while in the laboratory or studio we can create and combine auditory and visual streams to create illusions of unity, the reason that these illusions work is that in everyday life synchrony is indeed a reliable cue to a common source. Indeed, perfect synchrony is unlikely to occur by chance and is difficult to fake (although recall the Marx Brothers mirror scene in *Duck Soup*). While onset and offset are perhaps the most salient manifestations of synchronized dynamics, we are likely to notice a whole range of other correlations (e.g., pitch-space on the piano keyboard or a correlation between the size of a spider and the loudness of the spectator's scream) as well as make assumptions about likely correlations based on prior experience (large objects resonate at lower frequencies). Some auditory-visual associations we encounter are more arbitrary than others. For example, the visual appearance of the doorbell has very little predictive value concerning the sound heard when it is pressed. In novel situations, synchrony or co-localization alone may enable us to recognize a relationship (or perceive an illusory one).

One of the approaches to algorithmic synesthesia—simultaneous visualization and sonification of the same data set—does guarantee the presence of correlations that we are likely to encounter when sound and vision emanate from one source. How transparent or cognitively accessible the relationship will be is a different question, and we elaborate on this in the third part of this chapter.

D. Arbitrary Associations or Common Trends?

One of the puzzling features of synesthesia (the psychological phenomenon or condition) is the seemingly arbitrary nature of synesthetic correspondences. Synesthetes rarely agree about the specific colors they may associate with, for example, the letter *H*, Wednesday, middle C on the piano, or the taste of pistachio ice cream. Obviously, once a certain correspondence has been learned, it is no longer arbitrary for the person who has acquired it. Indeed, synesthetes are consistent in the cross-modal correspondences they experience across very long time periods (Sagiv 2005). Consequently, for a given individual, the feeling of seeing, for example, the color purple may have as much explanatory power about the timbre they hear (e.g., a trumpet) as the experience of the sound itself. The visual aspects become a part of the auditory experience. Shanon (2003) took this one step further and claimed that synesthesia entails more than cross-activation of different senses—it actually involves a relaxation of the boundaries between them.[4]

Interestingly, we find that once an association has been formed, we may also find evidence for bidirectional interaction even though the phenomenology is usually unidirectional. For example, synesthetes commonly report seeing numbers in color, but they do not "see" numbers when viewing a colorful display. Nevertheless, we see that given a numerical task, performance (reaction time) can be influenced by stimulus color (Cohen Kadosh et al. 2005). This effect can be present in addition to the more common finding that numbers may interfere with color judgments (number stimuli may slow color-naming time if the synesthetic color they evoke is incongruent with the presented target color).

Synesthetic correspondences are not completely idiosyncratic. In fact, in recent years we find that some trends seen in synesthetes are also found in nonsynesthetes. Such common inclinations are more consistent with the idea that synesthesia utilizes universal mechanisms rather and possibly represent an exaggeration of normal function rather than a gross abnormality (Sagiv and Ward 2006).

E. Common Auditory-Visual Associations in Synesthetes and Nonsynesthete Observers

Many synesthetes experience synesthesia in response to an auditory stimulus. Quite commonly, this is due to the linguistic aspects of a stimulus (Day 2005, Simner 2007). Still, some individuals automatically visualize colors when they listen to music or, in some instances, when exposed to a wide range of environmental sounds. Anecdotal reports suggested that for such synesthetes, experiences may depend on different factors—pitch, timbre, loudness, intervals, and chords. When synesthetes are asked to match colors to sound, they tend to choose brighter color with higher-pitch sounds (Marks 1975). Interestingly, this trend is also seen in nonsynesthetes (Marks 1974, 1987, Hubbard and Ramachandran 2005). A recent study by Ward et al. (2006) demonstrated this in a group of ten synesthetes and ten nonsynesthete controls using the same stimuli and methodology. They used seventy tones with varying pitch and timbre and asked participants to choose the color that went best with the presented tone. The procedure was repeated on a different occasion to establish consistency. While synesthetes were much more consistent in their responses than the nonsynesthete group, both showed essentially similar patterns. Regardless of timbre, all subjects associated brighter colors with higher pitch. However, more saturated colors were chosen at around middle C. Furthermore, timbre influenced saturation, regardless of pitch—pure tones were typically associated with less-saturated colors (low chromaticity). Although only synesthetes automatically visualized colors when presented with the sounds, the fact that nonsynesthetes made similar choices when required to match colors with tones is consistent with the idea that synesthetes utilize universal mechanisms. Still, we need to explain why some individuals have full-fledged synesthetic experiences while in most of us synesthetic associations remain below the threshold of conscious perception, manifesting only in behavioral measures and similarity judgments.[5]

Mondloch and Maurer (2004) showed that associations between pitch and brightness and between pitch and object size are present in preschool children (30–36 months old). Like adults, they associated higher pitch with brighter, smaller visual stimuli, although the association with size was weaker than the association with brightness. The fact that such associations are present very early in life is consistent with the neonatal synesthesia hypothesis (for a review, see Maurer and Mondloch 2005).[6] Marks (1975) also described a relationship between loudness of sounds and the size of the photism (louder sounds are associated with larger visual images).

F. Shared Meaning, Synesthesia, Language, and Gestures

We are able to share meaning through both a symbolic verbal system and nonverbal means. Some neuroscientists believe that we developed this capability using the "mirror neuron system" (e.g., Rizzolatti and Arbib 1998). Mirror neurons were first identified in the monkey premotor cortex (including an area homologous to the human language area). These neurons fire both when the monkeys perform a specific action or when they watch others perform it, thus creating a link between observer and the observed, enabling us to share meaning (at the very least, intentions) by simulating others' behavior. Ramachandran and Hubbard (2001) further suggested that certain common cross-modal associations may have facilitated the evolution of language, with natural constraints in the mapping of sound into objects providing the basis for a protolanguage.

Auditory and visual signals coming from a variety of sources (not necessarily human) can communicate meaningful information about their provenance. Some of these signals give rise to learned auditory-visual associations. More abstract stimuli produced (e.g., by musicians) can convey meaning when the listener appreciates the physical effort required to produce that sound, even when the gesture is not visible (Smalley 1992). Furthermore, when a gesture is present, it does have an impact on the experience of the listener (Dahl and Friberg 2007); for example, perceived sound duration can be influenced by the performer's gestures (Schutz and Lipscomb 2007). For the performer, the body also becomes synesthetically/kinesthetically linked to the auditory and visual aspects of the performance. For the audience, a body frame of reference becomes more important when the spatial extent of the stimuli varies or in interactive multimedia (e.g., Ben-Tal 2008). We return to the issue of spatial extent, and it is discussed from a different bodily perspective in chapter 11, this volume.

2. Cognitive/Algorithmic Underpinnings

We noted that perception and cognition not only align visual and sonic information but also tolerate or gain from their asynchrony. Such effects are well known in film; the sound designer in particular spends much effort on using quite unrealistic sonic action to amplify a realistic impression, and the sound may involve disparate noises as well as phase-shifted temporal associations (Chion 1994, van Leeuwen 1999). Thus, algorithmic control can vary the extent of perceptual association between sonic and visual events. It is interesting that in the creative arts it is common to anticipate multiplicity of interpretation on the part of vusers (viewer-listeners). The disparities between synesthetes in the nature of the reproducible synesthetic associations they experience potentially offers routes by which algorithmic synesthesia could amplify such multiplicity of all vusers' interpretations: the cognitive and affective responses of a person for whom the algorithmic alignment is consistent with their (synesthetic or merely preferential) associations will probably be rather differently influenced from a person whose associations are distinct.

Conversely, in the field of sonification, which seeks to make large quantities or rapid streams of data cognizable by representing them in sound, so-called zero-order transcoding involves complete synchrony between what is being represented and its sonic output. But, just as stock marketeers follow moving averages, representing blocks of data rather than individual events, so a range of sonification mechanisms perturbs this temporal relationship (see final section of this chapter).

It is often argued (see chapter 23, this volume) that humans can engage in several different modes of listening. In some, the emphasis is on everyday detection of events for the purpose of interpreting them in terms of their origins in some physical gesture. In others, so-called musical listening, sound is assessed for its intrinsic character with little reference to whether any natural object or event could generate such a sound. In computer music, commonly the sounds are such as can only be made by digital sound processing, techniques invented by humans. So, for sonification, the key question is to what degree the sound-generating algorithm is transparent, that is, permits the listener to understand the nature of the data being sonified. By extension, in algorithmic synesthesia, multimedia artists have the opportunity to place their processes anywhere along a spectrum of transparency of relationship between sound and image (and performer/creator) and to lead the vuser to establish personal understanding of such consistencies in this relationship as the artist builds in. All of this has to be taken in the context that with artistic multimedia, just as with physical environmental inputs, we discriminate in attention and importance between different signals, and just as we stop attending to the air-conditioning noise, so we may perhaps treat some components of a multimedia event. We may even be exposed to an excess of incoming information, beyond our capacity to handle.

Nevertheless, it is important to consider that while listening and viewing or while walking around the physical environment, we may also be generating our own distinct internal data streams, some of which may be imagined. Mental imagery for music is very often multimodal (see Bailes 2007). For instance, when we imagine a favorite piece of music, we may experience a mix of auditory, visual, and even kinesthetic mental representations. The multimodal nature of mental imagery is of considerable interest with respect to synesthesia, again underlining our propensity for cross-modal cognition. We might ask whether synesthetes experience cross-modal impressions when imagining sound or imagining vision. Very little is known beyond anecdotal report. While truly synesthetic experience may be idiosyncratic (although consistent), it seems likely that the triggering of a visual image by sound or a sonic image by vision for nonsynesthetes is largely contingent on the learned associations between sound and vision encountered through perceptual experience (or conditioning; see Cook 1998, p. 36). Eitan and Granot (2006) asked experiment participants to imagine the motions of a human character in response to simple melodic stimuli that were either an intensification or an abatement of a particular parameter such as dynamics, pitch contour, and pitch interval. They found significant relationships between certain musical parameters (such as pitch contour) and the imagined motion described by the presumably nonsynesthete respondents, and not all the relationships were as one might predict.

Ward (2004) argued that there is evidence of emotionally mediated synesthesia, describing the case of a synesthete for whom emotionally charged names of familiar people and other word categories were associated with a higher incidence of synesthetic color than emotionally neutral names and words. Perhaps emotion, familiarity, and attention play a similar role in the strength of cross-modal associations by nonsynesthetes. Some researchers have focused on suggested links between cross-modal perception, the perception of affect and artistic creativity (Dailey et al. 1997): participants who were classified as the most creative reported stronger associations between colors, pure tones, vowels, and emotional terms.

Should findings from the psychology of cross-modal perception and synesthesia guide us in the creative processes of algorithmic synesthesia? It is inevitable that artists will take into consideration some cognitive and perceptual constraints on the listeners and viewers, but the degree to which one seeks to mimic or recreate a genuine synesthetic experience is a matter of taste. Either way, there is always room for creativity, even when adopting a set rules or constraints and the results may or may not be interesting in either framework. The substantial variability in the phenomenology of genuine synesthetic perception demonstrates this and in fact guarantees that the artist retains considerable levels of freedom. Ultimately, even the attempts of a synesthete to convey his or her particular experience to others using sound, image, or both must involve some choices regarding which aspects of an incredibly rich personal experience to represent digitally. The difficulty may be due to the sheer complexity or due to attentional limits of the synesthete as an introspective observer (during the course of our work with synesthetes, we have often heard individuals complaining that they do not have time to observe all the synesthetic properties they perceive).

The observation that synesthetes often find their synesthetic experience pleasurable and aesthetically pleasing suggests that re-creating it does have an aesthetic potential. However, we must keep in mind that some aspects of synesthetic imagery may be very predictable and unremarkable; for synesthetic imagery to be interesting, the inducing stimulus may itself need to be interesting or novel. One factor with algorithmic synesthesia, as implied here, is that synesthetes in the audience may feel unease if their own synesthetic experience disagrees with choices made by the artists (e.g., "The colors are all wrong!"), but of course they must be used to such experiences in their day-to-day environments.

Let us now summarize potential strategies for algorithmic synesthesia based on the principles of naturally occurring auditory-visual synesthetic perception and some of the issues that they raise.

- *Synchrony:* Synchronized temporal patterns may be the most powerful cue for unity at a basic perceptual level. Once achieved, this unity can alter perception and so permit certain cross-modal interactions (keeping in mind that unity may be achieved with some small temporal misalignment). This is in contrast to nonsynchronous synesthesia, which may be linked to the evoking stimulus semantically, whether thematically or though learned associations. Having said that, some synesthetic imagery may be linked to more global parameters of a piece and not be temporally aligned.

- *Onset times* are not the only parameters we could manipulate. Direction of motion and location in space could also be used. We could covary pitch, loudness, and timbre with size, shape, hue, saturation, and lightness. It may prove particularly useful to use mappings found in both synesthetes and nonsynesthetes, such as pitch-lightness, pitch-size, loudness-size, timbre-chromaticity.
- *Different variants of color music synesthesia* include colored pitch, colored timbre, colored intervals, colored chords, or more global parameters (tonality of a piece). It is conceivable that color schemes based on any of these parameters could prove interesting (albeit disconcerting for some genuine synesthetes).
- *Similarly, language-based synesthesia could occur at different levels*—letter (grapheme), phoneme, whole words, or concepts. These may become useful when text or speech are included. More generally, this suggests the possibility of synesthetic streams arising simultaneously at different levels of processing of the same stimulus, modulated by selective attention to one level or another. In turn, we could potentially manipulate the audience's attention to different levels (or streams) by specifically choosing to visualize or sonify these rather than others.
- *Familiarity and consistency:* True synesthetes are very sensitive to deviations from their synesthetic correspondences, which may remain constant for years. It is unlikely that such fixity of correspondence could be achieved with a nonsynesthetic audience, even if considerable time is provided to allow them to learn associations. Thus, for the general audience, a deliberate change of mapping between sounds and image could still possibly serve creative aspects of algorithmic synesthesia.
- *Familiarity and meaning:* Synesthetic experience is sometimes weak for highly unfamiliar stimuli. It "grows" as the potential inducers acquire meaning. Using abstract, unfamiliar material may make it harder to form cross-modal associations on one hand, but on the other hand easier to accept a given scheme.
- *Finally, we need to beware of sensory overload.* This often occurs for genuine synesthetes and may be a concern in algorithmic synesthesia as well.

3. Algorithmic Synesthesia as a Creative Possibility

Through metaphor and synesthesia (cross-modal thinking and the unity of the senses), we acquire the means of enriching our experience of the symbolic codes of specific art forms by engaging with subsidiary meaning through other modes of perception (Cunliffe 1994, p. 163). The previous sections make quite clear, in agreement with Cunliffe (1994), that by sharing computational process or data, one may be able to establish creative exchanges between different modalities, and these may be cognitively and emotively interesting. In this final section, we briefly discuss some of the generic creative possibilities just listed, noting that such work is now widespread (cf. Whitelaw 2004, Dean et al. 2006), but little is systematic, and none has been subjected to empirical study, so we are at the beginning of our potential application of this mode of creative work.

For creative and algorithmic processes in particular, it is useful to envisage a continuum (or several) in which one can operate. There is probably only one way to conceptualize a single continuum between the visual and the sonic stimuli that encompasses all the materials relevant to algorithmic synesthesia: visual ↔ text-image ↔ speech ↔ music ↔ sound. We note that other aspects of synesthesia, such as those involving gustation, would require different treatment and are not addressed here. We note again that graphemic stimuli seem to be the most commonly potent. Moving and colored images of words or letters are now quite often used in creative writing for intermedia (see, e.g., Smith 2005). However, it is not clear that anyone has tried to use these relationships in the systematic manner that might challenge synesthetes, and more pertinently, relatively few have tried to use them in the potential manner of systematic algorithmic synesthesia. The opportunity is obvious.

Some software already encourages sound creation originating in image. For example, Metasynth is a program that permits the conversion of image to sound (e.g., color to spatialization or intensity) and to a more limited extent the converse process, and it is now widely used by composers, particularly in the creation of materials with which they will later work more intensively during an acousmatic composition. Conversely, many programs provide a continuous representation of a sound flux, mapped in simple or complex ways. As discussed, color and sound have strong relationships, and a long artistic history of artistic works, including Messiaen's *Chronochromie*, has sought to exploit them.

We have created subcontinua, for example between noise and speech, as have many composers (see chapter 14, this volume, on speech in computer music). The issues of synchronicity and relative speed of change in sound and image streams are already important in intermedia work and most highly formalized in film sound design. As mentioned, a perceived unity between sound and image (which may require some precise temporal misalignment) can create desirable versimilitude, which can eventually create cross-modal interactions by which one component becomes an index for the other. However, such synchronies are also the source of potential flexibility for composition developing and exploiting intermedia synesthesia, as in the case of the influence of multiple rapid aural events on the perception of the number of visual events they accompany (Shams et al. 2000, 2002, Wilkie et al. forthcoming). As suggested, a range of mappings of sound and image (e.g., pitch and darkness of color, etc.) deserve systematic attention. We have begun the empirical counterpart to such creative investigations by showing that rhythmic parameters of repetition of a sound (creating metrical patterns, for example), or to a lesser degree spatial parameters (alternating sidedness of the sound field), do have an influence on viewer perception of flashes. Thus, such approaches may be powerful in creative intermedia with complex sound and image.

It is particularly interesting to consider the future issue of three-dimensional (3-D) synesthesia and its possible algorithmic counterpart. A synesthete may project color onto the evoking stimulus, his or her immediate peripersonal space (which is within reach), and only rarely onto more distant space. It may be imaged on some internal mental screen, not clearly localized in space, or onto a bodily reference frame (such as the hand, or a screen "inside" the forehead). It will be difficult but fascinating to establish and control counterparts of these different processes in algorithmic synesthesia. Perhaps the most feasible opportunity is that of creating algorithmic synesthesia in the immediate peripersonal space of the vuser. This effect might coexist with their ability to move around a 3-D sonic field (the future traversable "sonic garden" that is open to investigation, as discussed in chapter 13, this volume). In such a future environment, the vuser could move both to and from sonic components and bring with them components of their own peripersonal sonic and visual image.

We have discussed (Dean et al. 2006) some examples of computer sonic works that at least engage the field of algorithmic synesthesia, such as from Robin Fox, Nick Collins, David Worrall and Stuart Ramsden, PanSonic, Andrew Gadow, and ourselves. But, essentially these use various forms of transcoding of sound into image or vice versa without as yet deeply assessing their potential for algorithmic synesthesia. Thus, in conclusion we note that the composer of a new work in a relatively unfamiliar idiom, such as many within acousmatic music, always has to deal with the need for new listeners to construct their own associations as the piece develops. Music, more than many arts because of its abstraction and limited external referentiality, provides special opportunities for interpretive multiplicity on the part of listeners. Algorithmic synesthesia foregrounds the corresponding issue in multimedia creative work, amplifies this multiplicity of interpretive possibility, and offers bold opportunities for contemporary intermedia.

Acknowledgments: We thank Oded Ben-Tal for helpful discussions and comments on the manuscript.

NOTES

1. Even under conditions of sensory deprivation (i.e., complete removal of sensory input), we may begin hallucinating before long (see, e.g., Merabet et al. 2004).

2. Consistency is rather easy to test for some types of synesthetic experiences (e.g., colors) by requesting synesthetes to choose a color from a given set or simply naming it. Other aspects may be more challenging to measure or quantify. Note, however, that there is no a priori reason to believe that a synesthetic correspondence must be consistent. While

consistency is well established in most recent reports of (developmental) synesthesia, it not clear that this is the case in drug-induced synesthesia.

3. For examples of illusory synchrony, see Nicholas Cook's discussion of "The Rite of Spring" sequence from *Fantasia* (Cook 1998, chap. 5).

4. Werner (1930, cited in Merleau-Ponty 1962/2002, p. 266) has noted this earlier: "For the subject does not say only that he has the sensation both of a sound and a color; it is the sound itself that he sees where colors are formed." Admittedly, for a nonsynesthete it is very difficult to understand what this means. It is not clear to us either, although we accept that a genuine synesthetic experience (whether naturally occurring or chemically induced) can be quite profound and perhaps qualitatively different from artificial multimedia.

5. We only have tentative answers to this question, and debates concerning the cognitive, neural, and genetic factors are ongoing. While this is outside the scope of this chapter, for further discussion see Merleau-Ponty (1962), van Campen (2008), and Ward (2008).

6. According to the neonatal synesthesia hypothesis, all newborns experience synesthesia (or at the very least cross-sensory confusion) until about 4 months of age. Indeed, electroencephalographic studies in infants and animal studies showed cross-activation of sensory areas, suggestive of higher connectivity within the brain. Some of these connections are pruned during development, but synesthetes may retain a higher proportion. Preliminary brain imaging studies suggest that this could indeed be the case (Rouw and Scholte 2007). Nevertheless, disinhibition or unmasking of existing connections has been suggested as a possible mechanism for acquired synesthesia associated with visual loss (Jacobs et al. 1981, Armel and Ramachandran 1999).

BIBLIOGRAPHY

Alais, D., and D. Burr. 2004. The ventriloquist effect results from near-optimal bimodal integration. *Current Biology: CB* 14(3): 257–262.

Alais, D., and S. Carlile. 2005. Synchronizing to real events: Subjective audiovisual alignment scales with perceived auditory depth and speed of sound. *Proceedings of the National Academy of Sciences of the United States of America* 102(6): 2244–2247.

Armel, K. C., and V. S. Ramachandran. 1999. Acquired synesthesia in retinitis pigmentosa. *Neurocase* 5(4): 293–296.

Bailes, F. 2007. The prevalence and nature of imagined music in the everyday lives of music students. *Psychology of Music* 35(4): 555–570.

Baron-Cohen, S., M. A. Wyke, and C. Binnie. 1987. Hearing words and seeing colours: An experimental investigation of a case of synaesthesia. *Perception* 16(6): 761–767.

Ben-Tal, O. 2008. "Musical Toys: Interactive Audio for Non-Musicians." Paper presented at Proceedings of the 6th Linux Audio Conference, Feb 28–March 2, 2008, Cologne, Germany.

Betancourt. 2007. A taxonomy of abstract form using studies of synesthesia and hallucinations. *Leonardo* 40(1): 59–65.

Botvinick, M., and J. Cohen. 1998. Rubber hands "feel" touch that eyes see. *Nature* 391(6669): 756.

Calvert, G. A., E. T. Bullmore, M. J. Brammer, R. Campbell, S. C. Williams, P. K. McGuire, P. W. Woodruff, S. D. Iversen, and A. S. David. 1997. Activation of auditory cortex during silent lipreading. *Science (New York, N.Y.)* 276(5312): 593–596.

Chion, M. 1994. *Audio Vision.* New York: Columbia University Press.

Cohen Kadosh, R., N. Sagiv, D. E. Linden, L. C. Robertson, G. Elinger, and A. Henik. 2005. When blue is larger than red: Colors influence numerical cognition in synesthesia. *Journal of Cognitive Neuroscience* 17(11): 1766–1773.

Cook, N. 1998. *Analysing Musical Multimedia.* Oxford: Clarendon Press.

Cunliffe, L. 1994. Synaesthesia, arts education as cross-modal relationships rooted in cognitive repertoires. *Journal of Art and Design Education* 13(2): 163–172.

Cytowic, R. E. 2002. *Synaesthesia: A Union of the Senses.* 2nd ed. Cambridge: MIT Press.

Dahl, S., and A. Friberg. 2007. Visual perception of expressiveness in musicians' body movements. *Journal of New Music Research* 29: 225–233.

Dailey, A., C. Martindale, and J. Borkum. 1997. Creativity, synaesthesia, and physiognomic perception. *Creativity Research Journal* 10(1): 1–8.

Day, S. 2005. Some demographic and socio-cultural aspects of synesthesia. In *Synesthesia: Perspectives from Cognitive Neuroscience*, ed. L. C. Robertson and N. Sagiv. New York: Oxford University Press, pp. 11–33.

Dean, R. T., M. Whitelaw, H. Smith, and D. Worrall. 2006. The mirage of algorithmic synaesthesia: Some compositional mechanisms and research agendas in computer music and sonification. *Contemporary Music Review* 25(4): 311–327.

Eitan, Z., and R. Y. Granot. 2006. How music moves: Musical parameters and listeners' images of motion. *Music Perception* 23:221–247.

Galeyev, B. M. 1988. The Fire of "Prometheus": Music-kinetic art experiments in the USSR. *Leonardo* 21(4): 383–396.

Hubbard, E. M., and V. S. Ramachandran. 2005. Neurocognitive mechanisms of synesthesia. *Neuron* 48(3): 509–520.

Ione, A., and C. Tyler. 2003. Neurohistory and the arts. Was Kandinsky a synesthete? *Journal of the History of the Neurosciences* 12(2): 223–226.

Jacobs, L., A. Karpik, D. Bozian, and S. Gothgen. 1981. Auditory-visual synesthesia: Sound-induced photisms. *Archives of Neurology* 38(4): 211–216.

Lochnan, K. A. 2004. *Turner Whistler Monet.* London: Tate Gallery Publishing.

Macaluso, E., and J. Driver. 2005. Multisensory spatial interactions: A window onto functional integration in the human brain. *Trends in Neurosciences* 28(5): 264–271.

Marks, L. E. 1974. On associations of light and sound: The mediation of brightness, pitch, and loudness. *American Journal of Psychology* 87(1–2): 173–188.

Marks, L. E. 1975. On colored-hearing synesthesia: Cross-modal translations of sensory dimensions. *Psychological Bulletin* 82(3): 303–331.

Marks, L. E. 1987. On cross-modal similarity: Auditory—visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance* 13(3): 384–394.

Martino, G., and L. E. Marks. 2001. Synesthesia: Strong and weak. *Current Directions in Psychological Science* 10: 61–65.

Maurer, D., and C. J. Mondloch. 2005. Neonatal synesthesia: A re-evaluation. In *Synesthesia: Perspectives from Cognitive Neuroscience*, ed. L. C. Robertson and N. Sagiv. New York: Oxford University Press, pp. 193–213.

McGurk, H., and J. MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264(5588): 746–748.

Merabet, L. B., D. Maguire, A. Warde, K. Alterescu, R. Stickgold, and A. Pascual-Leone. 2004. Visual hallucinations during prolonged blindfolding in sighted subjects. *Journal of Neuro-ophthalmology: The Official Journal of the North American Neuro-ophthalmology Society* 24(2): 109–113.

Merleau-Ponty, Maurice. 1962. *Phenomenology of Perception.* New York: Routledge and Kegan Paul. Repr. London: Routledge, 2002.

Mondloch, C. J., and D. Maurer. 2004. Do small white balls squeak? Pitch-object correspondences in young children. *Cognitive, Affective and Behavioral Neuroscience* 4(2): 133–136.

Nelson, M. E. 2006. Mode, meaning, and synaesthesia in multimedia L2 writing. *Language Learning and Technology* 10(2): 56–76.

Ramachandran, V. S., and E. M. Hubbard. 2001. Synaesthesia—A window into perception, thought and language. *Journal of Consciousness Studies* 8: 3–34.

Rich, A. N., J. L. Bradshaw, and J. B. Mattingley. 2005. A systematic, large-scale study of synaesthesia: Implications for the role of early experience in lexical-colour associations. *Cognition* 98(1): 53–84.

Rizzolatti, G., and M. A. Arbib. 1998. Language within our grasp. *Trends in Neurosciences* 21(5): 188–194.

Robertson, L. C., and N. Sagiv. 2005. *Synesthesia: Perspectives from Cognitive Neuroscience.* New York: Oxford University Press.

Rouw, R., and H. S. Scholte. 2007. Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience* 10(6): 792–797.

Sadato, N., A. Pascual-Leone, J. Grafman, V. Ibanez, M. P. Deiber, G. Dold, and M. Hallett. 1996. Activation of the primary visual cortex by Braille reading in blind subjects. *Nature* 380(6574): 526–528.

Sagiv, N. 2005. Synesthesia in perspective. In *Synesthesia: Perspectives from Cognitive Neuroscience*, ed. L. C. Robertson and N. Sagiv. New York: Oxford University Press, pp. 3–10.

Sagiv, N., J. Heer, and L. Robertson. 2006. Does binding of synesthetic color to the evoking grapheme require attention? *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 42(2): 232–242.

Sagiv, N., and J. Ward. 2006. Crossmodal interactions: Lessons from synesthesia. *Progress in Brain Research* 155: 259–271.

Schutz, M., and S. Lipscomb. 2007. Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception* 36(6): 888–897.

Seitz, J. A. 2005. The neural, evolutionary, developmental, and bodily basis of metaphor. *New Ideas in Psychology* 23(2): 74–95.

Shams, L., Y. Kamitani, and S. Shimojo. 2000. What you see is what you hear. *Nature* 408: 788.

Shams, L., Y. Kamitani, and S. Shimojo. 2002. Visual illusion induced by sound. *Brain Research. Cognitive Brain Research* 14(1): 147–152.

Shanon, B. 2003. Three stories concerning synaesthesia—A commentary on Ramachandran and Hubbard. *Journal of Consciousness Studies* 10: 69–74.

Simner, J. 2007. Beyond perception: Synaesthesia as a psycholinguistic phenomenon. *Trends in Cognitive Sciences* 11(1): 23–29.

Simner, J., C. Mulvenna, N. Sagiv, E. Tsakanikos, S. A. Witherby, C. Fraser, K. Scott, and J. Ward. 2006. Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception* 35(8): 1024–1033.

Smalley, D. 1992. The listening imagination: Listening in the electroacoustic era. In *Companion to Contemporary Musical Thought: Volume 1*, ed. J. Paynter, T. Howell, R. Orton, and P. Seymour. London: Routledge, pp. 514–554.

Smith, H. 2005 *The Writing Experiment: Strategies for Innovative Creative Writing*. Sydney: Allen and Unwin.

Steen, C. J. 2001. Visions shared: A firsthand look into synesthesia and art. *Leonardo* 34: 203–208.

Treisman, A. 1996. The binding problem. *Current Opinion in Neurobiology* 6(2): 171–178.

van Campen, C. 2008. *The Hidden Sense. Synesthesia in Art and Science*, *Leonardo*. Cambridge: MIT Press.

van Leeuwen, T. 1999. *Speech, Music, Sound*. London: Macmillan.

Viola, B. 1995. *Reasons for Knocking at an Empty House: Writings 1973–1994*. London: Thames and Hudson.

Wagner, R. 1849. *Das Kunstwerk der Zukunft*. Leipzig: Otto Wigand.

Ward, J. 2004. Emotionally mediated synaesthesia. *Cognitive Neuropsychology* 21(7): 761–772.

Ward, J. 2008. *The Frog Who Croaked Blue*. Oxford: Routledge.

Ward, J., B. Huckstep, and E. Tsakanikos. 2006. Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 42(2): 264–280.

Ward, J., and N. Sagiv. 2007. Synaesthesia for finger counting and dice patterns: a case of higher synaesthesia? *Neurocase: Case Studies in Neuropsychology, Neuropsychiatry, and Behavioural Neurology* 13(2): 86–93.

Ward, J., and J. Simner. 2003. Lexical-gustatory synaesthesia: Linguistic and conceptual factors. *Cognition* 89(3): 237–261.

Whitelaw, M., *Metacreation: Art and Artificial Life*. 2004. Cambridge, Mass.: MIT Press.

Wilkie, S., C. Stevens, and R. T. Dean. 2008. Psycho-acoustic manipulation of the sound-induced illusory flash. In *Lecture Notes in Computer Science*. Berlin: Springer, pp. 223–234

Zeki, S. 1999. Art and the brain. *Journal of Consciousness Studies* 6(6–7): 76–96.