



Sensory translation between audition and vision

Charles Spence^{1,2} · Nicola Di Stefano³

Accepted: 16 July 2023 / Published online: 6 October 2023
© The Author(s) 2023

Abstract

Across the millennia, and across a range of disciplines, there has been a widespread desire to connect, or translate between, the senses in a manner that is meaningful, rather than arbitrary. Early examples were often inspired by the vivid, yet mostly idiosyncratic, crossmodal matches expressed by synaesthetes, often exploited for aesthetic purposes by writers, artists, and composers. A separate approach comes from those academic commentators who have attempted to translate between structurally similar dimensions of perceptual experience (such as pitch and colour). However, neither approach has succeeded in delivering consensually agreed crossmodal matches. As such, an alternative approach to sensory translation is needed. In this narrative historical review, focusing on the translation between audition and vision, we attempt to shed light on the topic by addressing the following three questions: (1) How is the topic of sensory translation related to synaesthesia, multisensory integration, and crossmodal associations? (2) Are there common processing mechanisms across the senses that can help to guarantee the success of sensory translation, or, rather, is mapping among the senses mediated by allegedly universal (e.g., amodal) stimulus dimensions? (3) Is the term ‘translation’ in the context of cross-sensory mappings used metaphorically or literally? Given the general mechanisms and concepts discussed throughout the review, the answers we come to regarding the nature of audio-visual translation are likely to apply to the translation between other perhaps less-frequently studied modality pairings as well.

Keywords Sensory translation · Colour music · Crossmodal correspondences · Perceptual similarity · Synaesthesia

Sensory translation

The possibility of conveying the information associated with one sensory input by means of another or, at the very least, of trying to express one sensory impression by means of a sensation that is normally associated with the stimulation of another modality, has long attracted the interest of both scholars and artists. For example, a wide variety of inventors, designers, and artists have tried to translate music into light, colours and perfumes, while many novelists and poets have shaped their own style through the use of cross-sensory, or synaesthetic, metaphors, that is, ways of describing one

kind of sensory experience by means of another presented in a different sensory domain. Here, exploiting the traditional meaning of translation, conceived as “a craft consisting in the attempt to replace a written message and/or statement in one language by the same message and/or statement in another language” (Newmark, 1981, p. 7; see also Aguiar & Queiroz, 2013), we propose to refer to this broad spectrum of phenomena with the umbrella term ‘[sensory translation](#)’ (see Table 1 for a definition of a number of the key concepts appearing in this review).

The idea of the possibility of sensory translation might sometimes be rooted in the speculative historical claim that, although the senses can be conceived of as different channels with which to gather information from the environment, they must share some general functioning mechanism as they are all expressions of the same human perceptual abilities (e.g., Aristotle, 1907; see also Marks, 1978). Indeed, go back to ancient alchemical texts, and one sometimes comes across tables explicitly linking specific sounds to particular colours (e.g., see Huidobro Moya, 2007). Interestingly, scholars have suggested that the ancient Maya and Mesoamerican

✉ Charles Spence
charles.spence@psy.ox.ac.uk

¹ Crossmodal Research Laboratory, University of Oxford, Oxford, UK

² Department of Experimental Psychology, New Radcliffe House, University of Oxford, Oxford OX2 6BW, UK

³ Institute of Cognitive Sciences and Technologies, National Research Council of Italy (CNR), Rome, Italy

Table 1 Definition of key concepts appearing in this review

Key Concept	Definition	Relevant Sources
Translation	The attempt to replace a written message and/or statement in one language by the same message and/or statement in another language	Newmark, 1981
Sensory translation	The attempt to convey the meaning associated with one sensory stimulus by means of a stimulus presented in a different sensory modality	This review
Crossmodal correspondences	Deliberate and consistent matching between perceptual stimuli, attributes, or dimensions from different sensory domains that are observed in normal perceivers (i.e., non-synaestheses)	Spence, 2011
Synaesthesia	Neurological condition in which specific inducing stimuli give rise to an additional concurrent experience in either the same or different sensory modality	Grossenbacher & Lovelace, 2001
Inducer	The stimulus that triggers the synaesthetic perception	Grossenbacher & Lovelace, 2001
Concurrent	The stimulus that is elicited by the perception of the inducer in synaesthesia	Grossenbacher & Lovelace, 2001
Synaesthetic metaphors	Linguistic metaphors that cross the senses (e.g., ‘dark melodies’)	Hunt, 2005; Marks, 1996; Shen & Gil, 2007; Williams, 1976; Winter, 2011
Amodal dimensions	Properties of perceived objects and events that are allegedly common across different senses	Bahrick, 2009; Marks, 1978
Prothetic dimensions	Quantitative perceptual continua that have a clear ‘more than’ and ‘less than’ end	Stevens, 1957
Methatetic dimensions	Perceptual dimensions that obey a structured organization without necessarily having a ‘more than’ or ‘less than’ end	Stevens, 1957
Unidirectional/bidirectional	Referred to perceptual or sensory phenomena that occur in only one privileged direction, i.e., from one sense to another (unidirectional), or else in both directions (bidirectional).	Gil & Shen, 2021; Zhou & Tse, 2022

peoples experienced the senses—at least smell, hearing and sight—“as linked in a near-synaesthetic fashion” (Houston & Toube, 2000, p. 261). In Western philosophy, early conceptualizations of the senses stressed the distinction between different modalities while, at the same time, also emphasizing the intimate link between them. For instance, Aristotle presented the *sensus communis* as a putative psychological function connecting sensory impressions gathered from the five senses and processing them as a whole (Aristotle, 1907; see Johnstone, 2021).

In recent years, the question of whether it is possible to meaningfully connect the senses (or, at the very least, a subset of sensory impressions) has taken on renewed interest amongst those scientists trying to develop more intuitive sensory substitution devices (to compensate for sensory loss, e.g., in blind or deaf people, see Abboud et al., 2014; Hamilton-Fletcher et al., 2016; Marks, 1983; Spence, 2018, as well as for [sensory] augmentation, e.g., Akinbiyi et al., 2006; Quek et al., 2015; Pinardi et al., 2023). Sensory translation also appears frequently in texts on semiotics (e.g., Gottlieb, 2018), and has been gaining traction within the literature on museum studies as well (e.g., Liao, 2018; Neather, 2008, 2012; see also Cecon, 2011).

Over the centuries, a number of inventors and artists have been interested in trying to translate music into harmonious light displays/concerts, as with Alexander Wallace

Rimington’s colour organ (Jewanski, 2010a; Moritz, 1997; Peacock, 1988; Plummer, 1915; Rimington, 1895; Schöffner, 1985; Sullivan, 1914; see also Bragdon, 1916, 1918; Hector, 1922).¹ Cornelius Drebbel (1572–1633) is known as the inventor of musical instruments, such as virginals or lutes that would automatically translate light into sound (Gouk, 2000; Wilkins, 1680). However, the precise functioning of these innovative devices is not entirely clear. According to Wilkins (1680, pp. 148–149) when the instrument was placed in the sun, it would start to emit harmonies that were pleasant and soft. When moved into the shade it would become silent. This effect was allegedly caused by the warmth of the sun which affected some moisture in the instrument and the density on the air in it thus making its strings vibrate. The existence of this and similar instruments demonstrate that the human mind has long been attracted by exploiting the allegedly natural connection between sound and light and

¹ See Castel (1725, 1735, 1740, 1751, 1926a, b) for earlier accounts of an ocular harpsichord (Franssen, 1991; Hankins, 1994; Schier, 1941). According to Evans (1973, p. 190), in the late 1500s, the learned could even hope to perfect a ‘perspective lute’ that linked colours to music tones in synaesthetic union. There was also Athanasius Kircher’s (1602–1680) suggestion that each musical sound corresponded systematically and objectively to a certain colour (Cytowic, 1993; Kircher, 1650).

hence the possibility of translating something of the meaning or feeling associated with one sensory input into a stimulus from a different sensory modality (see also Mather, 2015).²

Other creative individuals, meanwhile, have wanted to convert music into perfumed performances (e.g., see Piesse's, 1867, 1891, 'Gamut of Odors'; and see Spence, 2021, for a review), or to deliver taste concerts by means of specially-constructed flavour organs (e.g., see 'The taste organ', 1926).³ Amongst the recent attempts to translate between audition and olfaction, Chang Hee Lee's (2013) project "Essence in Space" is worth mentioning.⁴ In this case, the translation is fixed by an adapted keyboard in which each key is mechanically linked to a fragrance situated below the keyboard. As each key is pressed, a droplet of perfume is released and collected in a bottle. This process continues as each key is struck, resulting in a mixture of different perfume droplets being collected. At the end of the 'performance', a unique blend of perfume has been created based on the olfactory conversion of musical "ingredients". In all such cases, those involved would appear to believe that there was a possibility of meaningfully connecting one sensory impression to another. Rarely, if ever, does one come across people proposing arbitrary cross-sensory mappings (i.e., under the assumption that any translation would be as good, or bad, as another). In a pedagogical context, Nijs et al. (2012) developed the Music Paint Machine, an interactive music system that translates movement and sound into colours, allowing a musician to create digital painting by playing an acoustic musical instrument and by moving on a coloured pressure mat. The system was conceived as a learning tool mainly aimed at the development of musical creativity and at strengthening the relationship between the musician and their instrument.

Around the start of the 20th century, many writers, including novelists and poets such as Charles Baudelaire (Anderson, 1980; Baudelaire, 1857, 1954) and Emily Dickinson, experimented with synaesthetic, or cross-sensory, metaphors in their work (e.g., Gibson, 1969; Harrison, 2001; von Erhardt-Siebold, 1932). In such cases, the artist's intuitive attempts to express one kind of sensory experience by means of another were often based on their own synaesthetic experiences (e.g., Di Stefano et al., 2022a; Marks, 1978). Intriguingly, however, such synaesthetic metaphors are typically unidirectional⁵

(Shen & Cohen, 1988; cf. Zhou & Tse, 2022), just as for synaesthesia proper (Deroy & Spence, 2013).

Over the last quarter of a century or so, a growing number of marketers and advertiser have become increasingly interested in the question of whether they can communicate (more) effectively with their customers by means of synaesthetic marketing (e.g., see Bolognesi & Strik Lievers, 2018; Crisinel & Spence, 2012a; Dunne, 2014; Kiefer, 2017; Meehan et al., 1998; Nelson & Hitchon, 1995, 1999; Spence, 2012). The latter can presumably be considered as yet another kind of sensory translation. Similarly, contemporary wine writers have been known to resort to a range of synaesthetic metaphors when trying to express the experience of flavour by means of musical analogies (Caballero, 2009).⁶ Intriguingly, if somewhat obscurely, Paradis and Eeg-Olofsson (2013, p. 22) have suggested that when terms such as 'sharp', 'soft', 'lemon', and 'cherry' are used to describe a wine's perceptual qualities, they should not be considered to be "polysemous synesthetic metaphors, but monosemous synesthetic metonymizations, more precisely zone activations." In this passage, it would seem that the authors would like to point out that the perception of such crossmodal similarities suggests that transitions across sensory domains in human language and understanding are not idiosyncratic, and tend to maintain the original semantic meaning unaltered (i.e., it is monosemous) rather than extending the literal meaning metaphorically to other domains (see also Rakova, 2003). Regardless, these phenomena fit with a growing variety of multisensory experiences, such as tasting events in which wine and music are deliberately paired (see Spence & Wang, 2015a, b, c, for reviews; and Spence, 2019a, for a recent review). Indeed, the topic of synaesthetic design has become increasingly popular in recent years (Haverkamp, 2014).

As will become clear, the concepts of synaesthesia, synaesthetic metaphor, and crossmodal associations play a key role throughout this review. While these concepts are closely linked, and have been often wrongly conceived as referring to the same phenomenon (i.e., as synonyms), they are also importantly different (as stressed by Cazeaux, 2002).⁷

² It might be interesting to note that, at that time, even most traditional instruments, such as the piano or the strings, had not reached their actual contemporary form yet.

³ Think here also of 'herbal capriccios' from Aldous Huxley's (1932) novel *Brave New World* (cf. Huysmans, 1884/1926).

⁴ See <https://www.changheele.com/essence-in-space.html>

⁵ Here, and throughout the paper, the adjective 'unidirectional' is attributed to a perceptual or sensory phenomenon that occurs from one sensory modality to another but not in the reverse direction. By contrast, the term 'bidirectional' refers to those perceptual phenomena that can be experienced both from one sense to another and vice versa.

⁶ Just take the following from British wine writer Hugh Johnson (2005): "I have tasted first-attempt Chardonnays that were like Dizzy Gillespie's solos: all over the place. And the colour of his trumpet, too. On the other hand a Stony Hill Chardonnay recently had the subtle harmonies and lilting vitality of Bix Beiderbecke. Robert Mondavi's Reserve Cabernets are Duke Ellington numbers: massed talent in full cry. Benny Goodman is a Riesling from Joseph Phelps, Louis Martini's wines have the charm and good manners of Glenn Miller. Joe Heitz, though, is surely Armstrong at the Sunset Café; virtuoso, perverse and glorious" (p. 253).

⁷ Waterworth (1997) introduces the term 'synaesthetic media', writing that: "The author posits that computer-based multimedia should be considered 'synaesthetic media' due to their capacity to translate between modes of sensory perception" (p. 327).

Synaesthesia refers to a rare neurological condition in which specific inducing stimuli give rise to an additional idiosyncratic concurrent experience in either the same or different sensory modality (Grossenbacher & Lovelace, 2001). Several different theories have been put forward to explain the existence of synaesthesia (see Simner & Hubbard, 2013, for a review). Here, we will be particularly interested in those cases of synaesthesia in which an inducer in one sensory modality gives rise to a concurrent in a different sensory modality, such as when hearing musical sounds elicits coloured concurrents (e.g., MacDougal, 1898).

The term ‘synaesthetic metaphor’ is typically used to refer to those linguistic metaphors that cross the senses, in expressions such as ‘a sharp-tasting cheese’, or ‘he is fond of wearing loud red trousers’. Such cross-sensory metaphor has been labelled synaesthetic because the cross-sensory expressions are similar, at least according to certain commentators, to the unusual inducer-concurrent mappings experienced by those synaesthetes who experience crossmodal forms of the condition. Some researchers have labelled such expressions as a kind of verbal synaesthesia (e.g., Popova, 2005). Finally, crossmodal associations, often referred to as crossmodal correspondences (Spence, 2011), are also often surprising to people when first hearing about them, just like synaesthesia. Crossmodal correspondences have been defined as the tendency for a sensory feature, attribute, or dimension in one sensory modality, either physically present, or merely imagined, to be matched (or associated) with a sensory feature, attribute, or dimension in another modality (Spence, 2011). Unlike synaesthesia, which is, by definition, idiosyncratic in terms of the inducer-concurrent mapping (Deroy & Spence, 2013; Grossenbacher & Lovelace, 2001), crossmodal correspondences tend to be consensual (see also Sun et al., 2018). The majority of people will, for example, associate round shapes with sweetness, and angular shapes with sour-tasting foods (Spence & Deroy, 2012). Furthermore, unlike is so often the case for synaesthetic metaphors (which typically only work, or really make sense, in one direction; see Chan et al., 2013; Gil & Shen, 2021; Zhou & Tse, 2022), the correspondences operate bidirectionally in the majority of cases (see Deroy & Spence, 2013; see also Smith, 1987, p. 94). So, while it makes sense to call cheese ‘sharp’, describing sharpness as a cheese really doesn’t work. By contrast, people are as happy to agree that sweet tastes are round (i.e., rather than sharp) as they are to match roundness with sweetness.

However, complicating matters somewhat, synaesthetes experience crossmodal correspondences (in this, they are no different from non-synaesthetes). On occasion, the inducer-concurrent mapping experienced by the synaesthete may coincide with the crossmodal correspondences experienced by the population at large. Intriguingly, while both synaesthesia and synaesthetic metaphor are typically considered to be unidirectional phenomena, the crossmodal correspondences that

have been documented to date are mostly bidirectional, thus highlighting a potentially important distinction between these superficially similar empirical phenomena (see also Deroy & Spence, 2013, for a summary of similarities and differences between crossmodal correspondences and synaesthesia).

When the diverse literatures are taken together, it soon becomes clear how widespread the interest in sensory mapping/translation really is (cf. Misdariis et al., 2021; Trotta et al., 2020, for a couple of other recent examples). Beyond its relevance to artistic performance or marketing/design contexts, we suggest that such a widespread interest might be rooted in the nature of the topic itself, which raises a number of intriguing (not to mention challenging) issues (see Daniels et al., 2010; Daniels & Naumann 2015). In this narrative historical review (see Ferrari, 2015; Furley & Goldschmied, 2021, on the strengths of narrative-style reviews), we attempt to shed some light on the topic by answering the following questions: (1) How is the topic of sensory translation related to synaesthesia, multisensory integration, and crossmodal associations? (see “[Synaesthetic Translations](#)”); (2) Are there common processing mechanisms across the senses that guarantee the success of sensory translation or, rather, is the mapping between the senses mediated by allegedly universal (e.g., amodal) stimulus dimensions? Answering these questions (see “[Putative Mechanisms Underlying Sensory Translation](#)” and “[Problems for Any Attempt to Translate Directly, or ‘Literally, Between the Senses](#)”), in turn, allows us to provide an answer to the following, more general, question as well: (3) Is the term ‘translation’ in the context of cross-sensory mappings used metaphorically or literally? The hope is that addressing these questions will help us to understand why it is that sensory translation has been considered, across the millennia, as a source of literary or poetic inspiration, a matter of philosophical reflection, and a research question worthy of serious empirical investigation.

The primary focus of the remainder of this review will be on the translation between audition and vision. In part, this simply reflects the fact that there is far more research on this pair of sensory modalities than for any other and, in part, the narrowing of focus was needed in order to prevent the review from becoming overly long. Nevertheless, the answers we come to regarding the nature of sensory translation between this frequently-studied pair of sensory modalities likely extend to the translation between other perhaps less frequently studied/mentioned modality pairings as well.

Synaesthetic translations

Synaesthetic perceptions

Historically, the most common source of inspiration for those wanting to express auditory(/visual) sensations by

means of visual(/auditory) ones has been the vivid, yet typically idiosyncratic, concurrents experienced by those individuals with synaesthesia. In fact, ‘coloured hearing’ turns out to be one of the most commonly-mentioned forms of synaesthesia, and often appeared in the scientific literature in the decades around 1900 (e.g., Argelander, 1927; Dauriac, 1902; de Parville, 1883; English, 1923; Flournoy, 1893; Ginsberg, 1923; Jewanski et al., 2009, 2011, 2020; Suarez de Mendoza, 1890; Underwood, 1893; Zigler, 1930). Both pitch (defined as a perceptual property of sounds that allows for their ordering on a frequency-related scale; see Zwicker & Fastl, 2013, p. 111) and timbre (also known as tone colour, or tone quality, from psychoacoustics, refers to the perceived sound quality of a musical note, sound or tone, see McAdams, 2019) appear to be salient auditory features (i.e., sensory inducers) driving the various coloured musical concurrents that have been reported in the literature (e.g., Curwen, 2018; Itoh et al., 2017; Marks, 1975). At the same time, however, it has also been acknowledged that there may be a strong visual mental imagery component to many coloured responses to music (Ahsen, 1997; Karwoski et al., 1942; Mudge, 1920; see also Mills et al., 2003; Nanay, 2018, 2020; Spence & Deroy, 2013). Referring to the evoked sensory experience in such cases in terms of ‘mental imagery’, rather than describing it as a synaesthetic concurrent, helps to draw attention to the fact that the qualities of the latter, such as the evoked colour coming or going, or else fading during a musical performance (see MacDougall, 1898; Riggs & Karwoski, 1934; Underwood, 1893), are not typical of synaesthesia as it tends to be conceptualized nowadays (see Spence & Deroy, 2013; though see also Nanay, 2018, 2020).

Many of the artists and composers who were interested in ‘colour music’ (e.g., Klein Cornwall-Clyne, 1937; Zilcher, 1987; see also Alves, 2005; Galeyev, 1976, 2003; McKellar, 1972, 1997; and Spence & Di Stefano, 2022b, for a recent review), such as, for example, Kandinsky (1977) and Scriabin, purportedly based their works on, or at the very least were inspired by, their own synaesthesia (e.g., Denham, 2017; Galeyev & Vanechkina, 2001; Harrison, 2001; Ione & Tyler, 2003, 2004; Kandinsky, 1977; Myers, 1911, 1914; Peacock, 1985; Witztum & Lerner, 2016; see Spence, 2020b, for a review). For example, Kandinsky (1977) suggested that the sound of the trumpet is scarlet (see Ione & Tyler, 2003, 2004; Just, 2017; though see also O’Regan, 2011). Kandinsky (1977) referred to a number of specific colour–sound mappings in his writings. However, it is often unclear whether the examples provided were based on the artist’s own synaesthesia, or else are perhaps better considered as examples of emotionally-mediated crossmodal correspondences (and hence might perhaps be expected to be experienced by us all; see Spence, 2020a, for a review), thus raising issues concerning how individual differences might affect the translation between the senses.

Something of a similar challenge faces those interested in trying to understand more about the idiosyncratic cross-modal mappings that have been suggested by synaesthetic Russian artists—namely, the composers Rimsky-Korsakov (who reported ‘seeing’ music in the key of A-major as yellow; Myers, 1911), and Scriabin (Galeyev & Vanechkina, 2001; Myers, 1914). Once again, though, it has long been the subject of debate as to what exactly the relationship, if any, was between Scriabin’s personal repertoire of idiosyncratic audiovisual inducer-concurrent mappings, and those chosen for his colour circle/score/luce (see Galeyev & Vanechkina, 2001; Triarhou, 2016). At this point, it is worth stressing that there is no real “translation” between two perceptual (and actually perceived) stimuli in the case of synaesthesia. More properly, with respect to the specific stimuli that are being linked (or associated), synaesthesia seems to be more related to synonymy, or identity, rather than necessarily translation since, for a synaesthete, the sound of the trumpet and the colour scarlet are simply part of one and the *same* perceptual experience (that is, the inducer is always co-experienced with the concurrent independent of the actual desire, or intention, of the perceiver).

Going deeper into the concept of translation, synaesthesia can perhaps be considered as a case of literal translation (Newmark, 1981), albeit one that is legitimated by “private” rules. That is, given the idiosyncratic nature of the synaesthete’s inducer-concurrent mappings, the concurrent can be seen as a faithful translation of, or synonymous with, the inducer, but it is faithful only for the individual synaesthete. By contrast, sensory translation can be considered more as a translation based on idiomatic and allegedly more universal criteria which can be shared across, or accessed by, normal perceivers. This means that synaesthetes cannot experience, for instance, the sound of the trumpet without, at the same time, also being aware of the colour scarlet, just as an English speaker cannot conceive of the terms ‘happy’ and ‘joyful’ as being unrelated. By contrast, nonsynaesthetes can experience roundness as unrelated to sweetness, just as an English speaker can conceive the term happy without necessarily thinking at the Italian translation “felice”.

Moreover, while it has been suggested that the inducer-concurrent mappings experienced by synaesthetes typically tend to be appreciated by non-synaesthetes (e.g., see Ward et al., 2008), it is important to note that this has no necessary implications for the question of whether the inducer can be (even metaphorically) considered as a ‘translation’ of the concurrent in the case of synaesthesia. According to O’Malley (1957, p. 393), synaesthesia might imply some loss of perceptual differentiation and discrimination between the inducer and concurrent, thus ruling out the possibility of looking for, or establishing, a comparison criterion underlying the association. As such, and especially because of the idiosyncratic nature of the inducer-concurrent mapping,

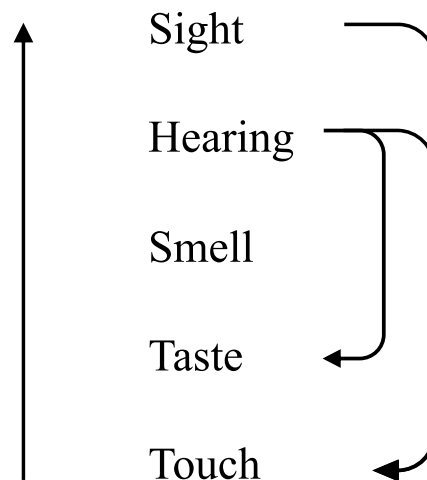
many researchers have been driven to search for alternative approaches to sensory translation that do not rely on the synaesthete's own idiosyncratic mappings from one sense to another. In the next section (“[Putative Mechanisms Underlying Sensory Translation](#)”), we examine whether the idea that the structural relationships between stimulus dimensions can help to provide a more robust means of translating between the senses, at least for a subset of stimulus dimensions.

Synaesthetic metaphors

Synaesthetic metaphors are expressions that “transfer one sense to another” (De Ullmann, 1945, p. 813; see also Shen, 1997) by establishing a relation between elements that are apparently semantically incompatible. For example, saying that a melody is sweet implies attributing a quality of taste to sounds, which do not manifest themselves through taste attributes. When it comes to the use of synaesthetic metaphors, there is an initial question as to which comparison modality people will intuitively gravitate toward and, thereafter, or perhaps as part of one and the same decision, which dimensions of sensory experience they deem it most appropriate to compare (Cazeaux, 2002; Motoki et al., 2020; Motoki & Velasco, 2021). Put more concretely, if one were to try and express the sound of the trumpet crossmodally, people might spontaneously reach for a visual metaphor (such as the colour scarlet) or perhaps a taste descriptor, an aroma, or perhaps a tactile texture instead (Shibuya et al., 2007). Over the years, a number of researchers have analyzed the patterns of ‘synaesthetic metaphor’ in both literature and poetry (see Hunt, 2005; Marks, 1996; Shen & Gil, 2007; Williams, 1976). Their findings provide insights concerning the modalities of expression (in terms of the direction) that occur most frequently in literary texts (see also Fishman, 2022).⁸

Many scholars (e.g., see Cytowic, 1989a, b, 1993; Day, 1996; Tsur, 1992; Ullman, 1957) have suggested that the perceptual modalities are organized along a scale ranging from the ‘highest’ modality—sight—followed by sound, smell (olfaction), taste (or gustation)—through to the ‘lowest’ sense—namely, touch (see also Houston & Taube, 2000, for a similar

⁸ Indeed, it has been suggested that the more successful literary metaphors may be those that are more immediate, with Downey (1912) suggesting, “It may be stated as a principle of interpretation that if a given transfer of sense-qualities is found pleasing only when the reader makes the suggested translation easily and spontaneously, there is some evidence that the writer himself used the analogy spontaneously rather than reflectively unless the expression be purely conventional” (p. 497). According to O’Malley (1957), “literary synaesthesia may imply a sort of introspection into basic processes of experience, a mirroring or echoing of the mind and senses to themselves in the act of apprehending phenomena” (p. 392). To be clear, synaesthetic metaphors are linguistic expressions in which a concept from one sensory domain is described in terms of another sensory domain (O’Malley, 1957, p. 391).



Shen and Aisenman (2008)
Strik-Lievers and Huang (2016)

Day (1996)

Fig. 1 Hierarchical organization of the senses based on their purported contribution to conceptualization and mental activity, ranking from the ‘highest’ modality—sight—to the ‘lowest’ sense—namely, touch (e.g., see Cytowic, 1989a, b, 1993; Day, 1996; Müller et al., 2022; Tsur, 1992; Ullman, 1957; Williams, 1976). Arrows indicate the most common direction of the occurrence of synaesthetic metaphors that are found in literary sources (Day, 1996), nonliterary texts (Strik-Lievers & Huang, 2016), and in both sources (Shen and Aisenman, 2008). [Reproduced from Di Stefano et al., 2022a]

hierarchy in ancient non-Western populations). According to Shen and Aisenman (2008), who reviewed a large corpus of literary and non-literary sources from different languages and cultures, synaesthetic metaphor (sometimes referred to as poetic metaphor, Marks, 1982a, b; linguistic synaesthesia, O’Malley, 1957; Sliz, 1942; or multisensory metaphor, Forceville, 2006) exhibit a robust, universal, tendency to use the ‘lower-to-higher’ structure more frequently than the inverse one (see Fig. 1). According to Conceptual Metaphor Theory (Lakoff & Johnson, 1980; see Landau et al., 2010, for a review), concrete terms provide the scaffold for more abstract terms/concepts. One might thus wonder if this concrete to abstract continuum can simply be extended to the case of synaesthetic metaphors (though see Paradis & Eeg-Olofsson, 2013, for an alternative viewpoint).⁹

⁹ According to Paradis and Eeg-Olofsson (2013, p. 37): “We call into question the metaphorical approach to words denoting sensory perceptions offered by Shen (1997) and Shen and Gadir (2009) which states that conceptual mappings from lower and more accessible concepts onto higher and less accessible are more natural than the other way round. They define conceptual accessibility through the directness of contact between the type of sensory perception and the object of experience. This explanation is dubious in many ways. Firstly, the notions of lower and higher modalities are not defined or agreed upon in the literature. In what sense is touch more accessible than smell and taste and why is vision not the most accessible modality?”

Many scholars have, though, wanted to argue against the existence of such a hierarchy amongst the senses. Perception is by its very nature multisensory (e.g., Spence & Bayne, 2015), typically involving the simultaneous stimulation of several senses with different sources of sensory information potentially overlapping. The encoding of such multisensoriality in language might also affect the quality of information that is conveyed or else the exclusion of certain aspects of the overall sensory experience (Winter, 2019). In this regard, language may represent something of a “biased” field in which to investigate sensory perception, as the mediation of cognitive structures might serve to reduce the complexity of sensory perception (see also Fishman, 2022; Gil & Shen, 2021; Pedović & Stosić, 2018; Simner et al., 2010). Combining neuropsychological evidence with the literature from the fields of cognitive and perceptual psychology, Winter (2019) argued against the very idea of synaesthetic metaphors, suggesting that the latter are “neither synaesthetic nor metaphorical”, rather reflecting the way language and perception are related and how sensory content is encoded in the lexicon of human languages.

Moreover, if metaphor can be used to explain certain associations, e.g., those between vision and audition, it may be somewhat harder to do for other sensory domains, such as olfaction, due to the apparent difficulty of establishing the underlying basis for the metaphor. For example, the results of a crossmodal matching task reported by Belkin et al. (1997) revealed the existence of a correlation (or correspondence) between auditory pitch and a range of olfactory stimuli. However, the exact nature of the underlying feature(s) responsible for this crossmodal alignment remains unclear. The authors hypothesized that participants might have based their matches on olfactory dimensions expressed in semantic terms, such as dull-aromatic (Wender, 1968), heavy-light, bright-dark, or hard-soft (Klutky, 1990).¹⁰ At the same time, however, Belkin et al. also note that if these associations were to have been based on metaphors then the latter were, at the very least, elusive both to the participants and to the experimenters who were studying them (see also Dubois, 2007; Juhasz, 1926, cited in Hartshorne, 1934, p. 240; Pomp et al., 2018).

Putative mechanisms underlying sensory translation

Structural mapping

Sir Ernst Gombrich (1960), the famous art historian, once suggested that we should focus our attention on the structural

relationships within the sensory systems rather than focusing on the similarity of specific elements when considering the nature of crossmodal associations. He suggested that “the problem of synesthetic equivalences will cease to look embarrassingly arbitrary and subjective if we fix our attention not on likeness of elements but on structural relationships within a scale or matrix” (p. 314). Gombrich was seemingly referring to the way in which stimulus dimensions are organized within each sensory modality. In this respect, it might be worth going back to Stevens’ (1957) early distinction between ‘metathetic’ and ‘prothetic’ sensory dimensions, with prothetic dimensions consisting of quantitative perceptual continua that have a clear ‘more than’ and ‘less than’ end. Examples of such prothetic sensory dimensions include loudness, brightness, lightness, heaviness, duration, and roughness. Metathetic dimensions, by contrast, tend to obey a well-structured organization without necessarily having a ‘more than’ or ‘less than’ end (see Table 1). Stevens classified the latter perceptual continua as ‘what kind’ or ‘where (position)’. For example, pitch is mentioned as a metathetic stimulus dimension, since a high-pitched tone is different in kind from a low-pitched tone, without necessarily being meaningfully related in a more than/less than way. Given this distinction, it seems reasonable to consider whether it is possible to translate between different prothetic dimensions by figuring out, or assuming, the relative position of the stimuli along their respective unisensory dimensions (cf. Cohen, 1934; Mellers & Birnbaum, 1982; Moul, 1930; Simpson et al., 1956, p. 100).

It might not be so surprising, therefore, that many of those who have attempted to translate between the senses (including artists, scientists, and writers) have chosen to match hue with pitch (Jewanski, 2010a; Sabaneev & Pring, 1929; Sabaneyev, 1911). This choice of modalities/dimensions may, in part, be related to the fact that we are all predominantly visual creatures (Hutmacher, 2019; see also Winter et al., 2018), and that, in vision, colours are particularly salient features of perceived objects. This, along with the fact that the structure of both auditory pitch and colour space can be represented metathetically (see Stevens, 1957) would also seem to have provided sufficient grounds for many to want to figure out some kind of meaningful relationship between this particular pair of sensory dimensions (see Spence & Di Stefano, 2022b, for a recent review). Importantly, according to Pridmore (1992), hue and pitch are also the only circular perceptual dimensions thus providing another reason for wanting to connect them (see Fig. 2).

Over the centuries, many different authorities have put forward their own solutions as to what might be the most appropriate translation between pitch and hue (e.g., see Field, 1835; Galejev & Vanechkina, 2001; Goethe, 1810/1840, c. 201–202, para. 748; Lavignac, 1899; Newton, 1704; see Table 2). This trend has continued over the

¹⁰ These, note, linking to the semantic differential approach (Osgood et al., 1957).

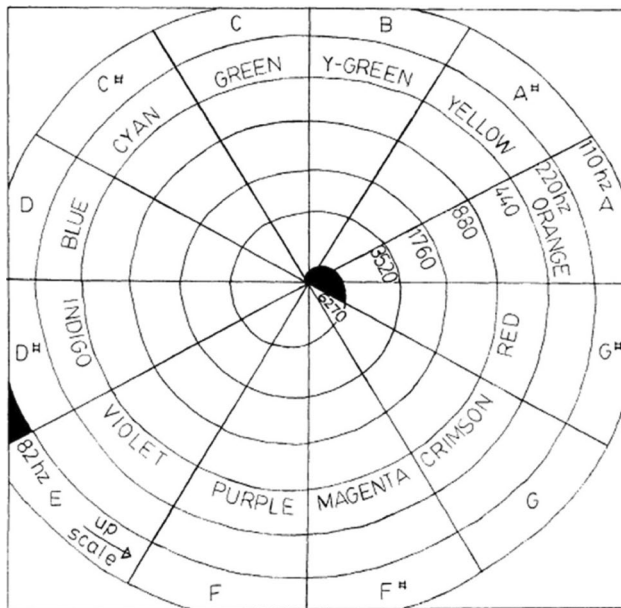


Fig. 2 Layout of Pridmore's (1984) final display panel. Notice how a given tone (e.g., C#) in all octaves (a musical octave is defined as a series of eight notes occupying the interval between [and including] two notes, the lower having half the frequency of vibration of the higher) is represented by a constant hue (e.g., cyan). Each octave is represented by a cycle, and each semitone (and its hue) by a sector (as presented in Pridmore, 1992)

last century, with a wide range of authorities, from architects to music scholars/teachers, and from psychologists to inventors, all putting forward their own (as it happens, somewhat idiosyncratic) cross-sensory mappings (e.g., Caivano, 1994; Davis, 1979; Galeyev & Vanechkina, 2001; Garner, 1978; Pridmore, 1992; Sebba, 1991; Wells, 1980). However, whenever experimental psychologists have attempted to demonstrate the influence of such crossmodal correspondences (e.g., between hue and pitch), using, for example, the speeded classification task, they have not been especially successful (see Bernstein et al., 1971; Melara, 1989), or else their results have been queried by psychologists on methodological grounds. So, for example, one of the main criticisms that has been levelled at Simpson et al.'s (1956) classic study, apparently showing a relationship between pitch and hue, is that the lightness of the visual stimuli may not have been controlled and hence any crossmodal mapping that was reported may actually reflect a lightness-pitch correspondence instead (e.g., Wicker, 1968).

In summary, given the fact that neither synaesthesia (see [Synaesthetic Translations](#)), nor the structural approach to capturing, or establishing, perceptually meaningful audiovisual associations works, it would seem appropriate to look for an alternative solution, or theoretical framework, with which to conceptualize the crossmodal translation/matching of pitch and hue.

Perceptual similarity

As O'Malley (1957) observed, the “metaphorical fusion of different sense data must always have carried intimations of a plane or phase of reality in which there was indeed a sharing, an interchange of properties” (p. 409), thus suggesting that sensory translation might well be rooted in those properties that are shared by different stimuli. Indeed, the existence of consistent crossmodal associations might be taken to reveal that the stimuli that are paired are, in some respect at least, similar (O'Malley, 1957, p. 392). Going back to the audiovisual domain, while the various structural attempts to match pitch with hue have seemingly failed to arrive at any consensus, it is noticeable how, when freed from the constraint of having to align dimensions of sensory experience due to their structural similarity, some authors have instead reached for timbre-hue crossmodal mappings instead. In fact, long before Kandinsky came out with his suggestions concerning the crossmodal association between the sound of a trumpet and the colour scarlet (see Kandinsky, 1977), both Locke (1690) and Leibniz (1704/1896) had already put forward essentially the same crossmodal mapping when considering the experience of a blind man. The composer Raff also reported that he perceived the colour of the sound of the trumpet to be scarlet (other people apparently report it to be bright red; Ortmann, 1933), while, for Kandinsky, the sound of the tuba was also red (see also Anikin & Johansson, 2019; Donnell-Kotrozo, 1978; Ginsberg, 1923, p. 589; Menouti et al., 2015; Reuter et al., 2018b). Other contemporary researchers, meanwhile, have chosen to study the correspondences between timbre and shape (see Adeli et al., 2014; Gurman et al., 2021) or harmony (namely, consonance and dissonance) and visual roughness (Giannos et al., 2021; see Di Stefano & Spence, 2022, for a review on multisensory roughness, and Di Stefano et al., 2022b for a review on consonance and dissonance) (Table 3).

Wicker (1968) conducted an intriguing pair of early studies designed to investigate the “intersensory dimensions in perceptual or cognitive space, *i.e.*, of dimensions which are significantly descriptive of sensory inputs from more than one modality” (Wicker, 1968, p. 178, italics in original). In a first experiment, Wicker's participants were presented with a range of 13 pure tones (300, 400, 500, 600, or 700 c.p.s.) of varying loudness (53–84 dB); they were also presented with 13 coloured Munsell colour squares (green, red, blue, and yellow) of varying brightness and saturation. The participants had to rate the similarity of all pairs of tones and thereafter to rate the similarity of all possible pairs of colour patches. They also had to rate every individual tone and colour patch in terms of 25 semantic differential adjective scales (cf. Moller et al., 2009), as well as to rate the similarity of the auditory and visual stimuli. In particular, they had to rate the similarity

Table 2 Correlation of colour and notes of the chromatic scale. Table highlighting the various crossmodal correspondences that have been proposed since Newton (1704). Newton’s correlation conforms to the seven-tone scale which he was probably familiar with. Castel’s correlations were made with the 12-tone chromatic scale, but as Wells (1980) notes, the hues fall in frequency as the tones rise in frequency.

The alignment is reversed for the following correlations. The scale attributed to E. G. Lind presents the pitch of tones (sound frequency, Hz) and the frequency of light (presented in parentheses in the Table as Terahertz (THz), for example red is 476 THz). Rimington’s (1895) crossmodal mapping from Peacock (1988). Adapted from Wells (1980, Table 1)

Note	Newton 1700	Castel 1720-1735	Finn 1881	Rimington 1893	Lind 1900	Maryon c. 1920
C	Red	Blue	Red	Deep red	259 Hz, red (476)	Red
C#		Sea green, blue-green	Vermillion	Crimson		Red-orange
D	Orange	Green, bright green	Orange	Orange-crimson	289 Hz, orange (511)	Orange
D#		Olive, yellow-green	Yellow			Orange-yellow
E-flat				Orange		
E	Yellow	Yellow	Yellow-green	Yellow	322 Hz, yellow (546)	Yellow
F	Green	Apricot, yellow-orange, aurora	Green	Yellow-green	342 Hz, green (588)	Yellow-green
F#		Orange	Blue-green	Green		Green
G	Blue	Red	Turquoise blue	Bluish green	385 Hz, blue (630)	Blue-green
G#		Crimson	Blue			Blue
A-flat				Blue-green		
A	Indigo	Violet	Indigo	Indigo	427 Hz, indigo (665)	Blue-violet
A#		Agate, blue-violet, light purple	Violet			Violet
B-flat				Deep blue		
B	Violet	Indigo	Purple	Violet	485 Hz, violet (721)	Violet-red

of every tone to every colour using a 9-point scale (Wicker, 1968, p. 180).

Multivariate scaling revealed two orthogonal alignments underlying the intersensory and cognitive space: pitch-brightness and loudness-contrast. The correspondences between these dimensions were established using multidimensional-scaling (MDS), semantic-differential (SD) scaling, and an intersensory transfer of training paradigm. At the same time, however, Wicker’s (1968) results failed to provide evidence of the existence of any crossmodal correspondence between pitch-saturation, loudness-brightness, and loudness-darkness. These findings would therefore appear to suggest that the mapping of colour to sound is likely to have been based on frequency (in terms of the low-high continuum) while ignoring octave repetition.¹¹

While it is commonly accepted in the academic literature that it may be possible, and meaningful, to make judgments concerning the perceptual similarity of pairs of stimuli presented within the same sensory modality

(Ekman, 1954; Ekman et al., 1964; Marks & Bornstein, 1987; Shepard, 1962, 1974; Tversky, 1977; von Helmholtz, 1878/1971), talking about the perceptual similarity of stimuli presented in different sensory modalities is more controversial, and has attracted a much more diverse range of theoretical responses (e.g., Di Stefano & Spence, 2023; O’Regan, 2011; Spence, 2022). For example, von Helmholtz refuted the very idea that perceptual similarity had any meaning across the senses when he wrote that “the distinctions among sensations which belong to different modalities, such as the differences among blue, warm, sweet, and high-pitched, are so fundamental as to exclude any possible transition from one modality to another and any relationship of greater or less similarity. . . . Comparisons are possible only within each modality” (von Helmholtz, 1878/1971, p. 77; though see Hartshorne, 1934).¹² Lawrence Marks (2011), by contrast, had the following to say concerning: “Perceptual similarities between and among sensory experiences in different modalities. Much

¹¹ Cf. Cytowic and Wood (1982), and Ries (1969), for a couple of other examples of participants being asked to pick colours to match a range of tones.

¹² Or, as Dr. S. Johnson (1751) put it: “Sound can resemble nothing but sound.”

Table 3 Summary of different kinds of audiovisual crossmodal mappings/translations based on different kinds of stimulus dimension

Visual dimension		Auditory dimension	
Brightness (Prothetic)	Hue (Circular & Metathetic)	Brightness (Prothetic)	Shape (No obvious organization)
Putatively amodal (e.g., Lewkowicz & Turkewitz, 1980; Walker-Andrews, 1994), and hence also structurally-based) Loudness-luminosity (Caivano, 1994); Can also be considered as prothetic correspondence (Spence et al., 2013); NB. Lightness & brightness may elicit different correspondences; (Marks et al., 1987)	Anikin & Johansson (2019) report loudness associated with saturation; Hamilton-Fletcher et al. (2017) correspondence between frequency & chroma; (NB. Colour sometimes confounded with lightness)	No evidence	No evidence
Pitch (Circular & Metathetic) See Wicker (1968); Polarity mapping/ correspondence between low vs. high pitch; white vs. black (Ludwig et al., 2011; Marks, 1987a; Melara, 1989); (NB. in polar correspondences, it tends to be the relative, rather than absolute pitch that matters, see Spence, 2019b)	Structural alignment of circular dimensions (Caivano, 1994; Field, 1835; Galeyev & Vanechkina, 2001; Goethe, 1840; Lavignac, 1899; Newton, 1704; Pridmore, 1992; Sebba, 1991; Wells, 1980). (NB. Pitch change may introduce loudness difference)	Turned into polar mapping between low vs. high pitch, and round vs. angular shape (Marks, 1987a; Parise & Spence, 2012; Wellek, 1927)	
Timbre (No obvious organizational principle)	Anikin & Johansson (2019); Donnell-Kotrozo (1978); Menouti et al. (2015); Mudge (1920); Reuter et al. (2018a)	No evidence	Adeli et al. (2014); Gurman et al. (2021); (NB. Smith & Sera, 1992, suggest that shape is a metathetic dimension)

as the color aqua is more similar phenomenologically to cerulean than to pink, the flavour of lime more similar to lemon than to banana, so too are low notes played on a bassoon or an organ more like dark colors such as brown or black than bright colors such as yellow or white, while the higher notes played on clavier or a flute resemble yellow or white more than brown or black” (p. 52). Elsewhere, Marks (1996) suggests that the best we can hope to achieve is what he once called ‘perceptual metaphors’—a term that can perhaps be taken to be synonymous with synaesthetic metaphor.

The above quotes thus demonstrate that eminent psychophysicists have seemingly taken diametrically opposed positions concerning the very possibility of grounding a meaningful translation between the senses on perceptual similarity. At the same time, however, more general concerns can be raised regarding perceptual similarity when the notion is evoked in the cross-sensory context to explain why it is that certain complex stimuli are associated with one another. For example, a funeral march might be associated with a weeping willow as they are both perceived to be sad. In such a case, sadness is the property shared by the paired stimuli. However, the fact that people tend to associate X to A more than to B, for instance, the song “Happy Birthday” to sunflowers more than to weeping willows, does not in-and-of-itself necessarily imply that X is perceptually similar to A. While audiovisual semantic congruency¹³ is established on the basis of the regular co-occurrence of the component auditory and visual stimuli, this does not have any necessary implications for the question of whether the component stimuli are themselves perceived as being in any way perceptually similar (Wegner-Clemens et al., 2022; see also Di Stefano & Spence, 2023).

Furthermore, demonstrating the existence of a statistically significant (or consensual) crossmodal correspondence between stimuli only shows that the pairing was the best of the options that were made available for participants at the time that they were asked (see Spence & Levitan, 2021, on this point). Thus, the mere fact that a robust crossmodal correspondence can be established between two stimuli does not mean that the stimuli are necessarily similar in some/any respect. For example, if English-speaking participants were to be exposed to the word ‘apple’ and to the images

of an apple and of a tree, and were asked to match the word with one of the two images, they would undoubtedly match the word to the image of the apple. Clearly, though, in this case, the word ‘apple’ is not more perceptually similar to the apple than it is to the tree. Note that according to embodied theories of language processing (e.g., Barsalou et al., 2012), the exposure to words is thought to activate perceptual representations that are consistent across individuals. However, this does not mean that the words are in any intuitive way similar to the representation they elicit, for instance the word “apple” to the apple, nor that any kind of similarity triggers the perceptual representation, which can probably be more easily explained in term of association.

Another theoretical issue would appear to weaken the idea of perceptual similarity based on shared phenomenological properties even further. The problem here is that a thing might be an ‘A-thing’ with respect to ‘A-ness’, and at the same time a ‘B-thing’ with respect to ‘B-ness’ (see Rodriguez-Pereyra, 2002, for an extensive discussion of this problem). To give an example, to help make this point more concrete, say “redness” is A-ness, while “being a fruit” is B-ness. A strawberry is both an A-thing and a B-thing. So, it is similar to all red things, say cherries and blood, but also to peaches and grapes. This suggests that grounding similarity relationships on shared phenomenological properties will likely make similarity a universal relationship (you will always find an A-ness according to which two different things are similar). Critics point to the fact that any two objects might share at least one phenomenological property and thus, as Goodman (1972) has argued, similarity would simply be a universal relation—namely, everything would be similar to everything else—and therefore claims regarding similarity would become somehow meaningless. Moreover, different properties count differently as far as perceptual similarity is concerned. For example, a strawberry is more similar to a blueberry than to blood, despite both strawberries and blood being red. Thus perceptual similarity would appear to depend on more than just a simple count of shared and unshared perceptual features or attributes; that is, it depends on emergent Gestalt properties (e.g., Palmer, 1989; Pomerantz et al., 1989; Di Stefano and Spence (2023) recently presented an account of perceptual similarity based on a two-dimensional space with associative strength on one axis, and cognitive penetrability on the other).

A further possibility we will investigate is simply that there is no similarity between perceptual dimensions across the senses, but rather that there exist a certain subset of perceptual dimensions that are amodal, or supramodal, meaning (at least according to certain commentators) that the same information is picked-up regardless of the sensory source (Walker-Andrews, 1994). O’Malley (1957, p. 392) talks of

¹³ The term ‘semantic congruency’ is somewhat confusingly used in different ways in at least two different literatures. Here, the term is used to refer to those situations where the different sensory properties of a source object, such as the sight and sound of a dog, are combined in semantic representations, as a result of prior co-occurrence (see Chen & Spence, 2011; Hsiao et al., 2012). Note that semantic congruency has no necessary relation to perceptual congruency, the latter being based on perceptual similarity. Consider here only the fact that while the image of the dog and the sound of its bark are semantically connected, they are not in any sense perceptually similar (see also Spence, 2022, on this distinction as it pertains to the chemical senses).

this as ‘intersense analogy’,¹⁴ going on to say that: “In literary discussions, intersense analogy and clinical synaesthesia are seldom distinguished, but it is important to stress their essentially different implications. The principal difference concerns the question of whether or not intersense comparisons or resemblances are accessible to normal, if heightened, experience. For clinical synesthesia, the question of resemblance is incidental; for intersense analogy, it is essential” (O’Malley, 1957, p. 393). Notice how here, O’Malley would appear to be drawing attention to the idiosyncratic nature of the inducer-concurrent mapping in synaesthesia.

Amodal properties

More than two millennia ago, Aristotle drew attention on the existence of features of the world that can be perceived in their own right by different senses: “For the perception of magnitude, figure, roughness, smoothness, and sharpness and bluntness, in solid bodies, is the common function of all the senses, and if not all, then at least the common function of sight and touch” (*De Sensu et Sensibili*, 442b in Aristotle (1908); see also *De anima*, 418a10–11, 19 in Aristotle (1907); see also Paterson, 2021, p. 33; Werner, 1934, p. 202). Based on such an almost synaesthetic view of perception, some researchers have wanted to suggest that amodal sensory dimensions might therefore provide a robust basis for connecting the senses. A number of developmental psychologists have argued that amodal stimulus dimensions provide a fundamental role in terms of helping to scaffold multisensory interactions in human development (e.g., Bahrick, 2010; Bahrick et al., 2004; Bahrick & Pickens, 1994; Gibson, 1969; Lickliter & Bahrick, 2012; Smith, 1987).¹⁵ Consider here only how people can recognize the same temporal pattern no matter whether the information is provided via hearing, touch, or vision (e.g., Marks, 1987a; see also Frings & Spence, 2010; Marks, 1987b), thus suggesting that such mechanisms that enable the processing of temporal patterns are not modality-specific, and hence that temporal pattern is likely to be an amodal stimulus property. According to Bahrick (2009): “Properties of objects and events such as temporal synchrony, rhythm, tempo, duration, intensity,

and co-location are common across auditory, visual, and proprioceptive stimulation” (p. 44; see Grahn, 2012; Huang et al., 2012; McAuley & Henry, 2010). Taken together, such results would appear to support Lewkowicz and Turkewitz’s (1980) early claim that rhythm is an amodal dimension. A similar argument has been made with respect to numerosity (see Gallace et al., 2007). It should, though, be noted that both numerosity and rhythm are linked to the organization of groups of stimuli and are not themselves sensory properties of the stimuli (i.e., in the way that stimulus intensity is, say).

Bahrick (2009) notes that amodal literally means “without” modality. However, she chooses to ignore this use of the term (e.g., as the term is used in the literature on the perceptual completion of occluded stimuli). Instead, the dimensions that Bahrick considers amodal consist of a mixture of those that are amodal by virtue of the fact that different senses sometimes pick up the same sensory information, such as vision and touch providing information about the size and shape of hand-held objects, and those that are amodal in virtue of the fact that the perceptual quality is somehow equivalent across different senses, as in the case of stimulus intensity. Note, though, that there is seemingly no explicit necessity for the relevant unisensory experiences nor for what is being picked-up from different senses to be phenomenally similar (Ernst & Banks, 2002). According to Bahrick (2009), “Amodal information includes changes along three basic parameters of stimulation—time, space, and intensity” (p. 44). Marks et al. (1987) also talk of the “perceptual, cross-modal equivalence with respect to intensity” (p. 5).¹⁶

One of the earliest studies on amodal dimensions was published by von Hornbostel (1931). Von Hornbostel’s hypothesis was that brightness represented a universal dimension of sensory experience. The small number of participants ($N = 3$) in his study had to match sounds of different pitches to points along a greyscale. They also cross-modally matched scents with grayscale values. Given the apparent transitivity between different crossmodal comparisons, von Hornbostel interpreted his results as demonstrating that the concept of ‘sensory brightness’ must be common to all of the senses. Were this, in fact, to be the case, one could easily imagine how simply matching the brightness of

¹⁴ O’Malley (1957) writes that “intersense analogy stands for the supposed comparability or resemblance of qualities which are perceived by different senses and may refer to single, tentative parallels like Aristotle’s or to detailed and systematic comparisons of the data of two or more sensory modes” (p. 392).

¹⁵ For example, just take the following suggestion from the developmental psychologist Linda Smith (1987) that “we have a lot to gain if we can get beyond mapping specific dimensions one to another and instead delineate the amodal dimensions” (p. 96). Clark (1970), meanwhile, argues that polar dimensions are fundamental to children’s development of relational concepts.

¹⁶ Paradis and Eeg-Olofsson (2013) write that “as pointed out by Rakova (2003, p. 49), similarities due to perceptual equivalence, rather than conceptual analogies have been extensively studied. Those studies confirm that the perception of cross-modal similarities is universal, systematic and present in early childhood. What these results suggest then is that transitions across sensory domains in human language and understanding are monosemous and syncretic rather than metaphorical and polysemous. Rakova (2003, p. 142) claims that one may think of concepts such as BRIGHT, SHARP, and COLD as primitive concepts spanning all domains of sensory experience. They may be thought of as neural configurations responsive to certain stimuli” (p. 37).

auditory and visual stimuli would provide a means of meaningfully (or consensually) translating between the senses.¹⁷

Even early researchers were not entirely convinced of the existence of amodal stimulus dimensions (see Cohen, 1934), arguing instead for a relative/relational judgment account (i.e., rather than necessarily a crossmodal perceptual mapping based on amodal properties; cf. Hartshorne, 1934). As Cohen (1934, p. 119) put it, and importantly for our concept of sensory translation (where two different stimuli are necessarily at stake), the stimuli in von Hornbostel's study were 'analogous' rather than 'identical'. Cohen tried to explain his reasoning as follows: "It would not be unreasonable then to suppose that cross-modality comparison should be based (physiologically, if not introspectively) upon relative positions within different 'absolute' scales. According to this view equation with respect to brightness of two experiences of different modalities would involve nothing more than the identity of relative positions upon two wholly independent scales." According to Marks et al. (1987, p. 34): "As a general rule, psychophysicists who study crossmodal matching have concerned themselves primarily with determining precise quantitative measures of intersensory equivalence; their purpose is usually to test theoretical predictions made from psychophysical functions (which relate judgments of sensory magnitudes to physical intensities) derived for individual continua like loudness and brightness." In other words, while a robust psychophysics of crossmodal matching is consistent with the existence of an underlying amodal dimension guiding people's choices, it certainly doesn't entail it.

Mellers and Birnbaum (1982) describe the distinction thus:

Two prominent theories of cross-modality matching are mapping theory and relation theory (Krantz, 1972; Shepard, 1974). According to mapping theory, psychological values of stimuli from different continua are mapped onto a common scale of sensation and can be compared directly. A cross-modality match is presumed to occur when equal strength sensations are elicited by stimuli on different continua. According to relation theory, relationships (e.g., ratios) between pairs of stimuli from different continua are compared. In physical measurement, a mass in grams cannot be compared with a length in centimeters but ratios of masses can be compared with ratios of length. By analogy, it may be possible to compare the ratio of the heaviness of two weights to the ratio of the loudness

of two tones, since the ratios of stimulus pairs are on a common scale. (p. 593)¹⁸

Later, Mellers and Birnbaum (1982, p. 600) go on to suggest that: "In cross-modality judgments, the scale values are influenced by the stimulus distribution: It appears that subjects compare the relative position of a stimulus in its distribution with the relative position of a stimulus of another modality to its distribution", going on to suggest that their results were consistent with a psychological relativity theory of crossmodality judgment.

The possibility that amodal concept(s) might exist is apparently linked to the existence of absolute correspondences,¹⁹ as Smith (1987) observed:

This suggestion of a trend from dichotomous, categorical treatments of continua to more relativistic ones ought not to be confused with the issue of absolute versus relative correspondences across dimensions. The notion of absolute correspondences between dimensions is that particular values on one dimension map onto particular values on another—for example, higher is not like brighter; rather, a specific pitch matches a specific brightness. As Marks et al. point out, though, there is little evidence for such absolute correspondences. (pp. 97–98)

Note here also that pitch-based crossmodal correspondences tend to be relative rather than absolute (see Spence, 2019b).

Marks et al. (1987) conclude that "in some fundamental sense, the similarities between pitch and brightness and between loudness and brightness are personal, internal, and subjective; they reside in perception per se and probably depend on common processes of neural coding" (p. 84). Note the strong claim here, albeit with multiple provisos, that similarity relations are perceptual in nature (see Di Stefano & Spence, 2023, for a discussion of the perceptual/cognitive nature of similarity). Nevertheless, the 'personal, internal, and subjective' element did not stop Marks (1987a) from trying to establish a robust psychophysics based on the crossmodal matching of the colour lightness of grey surfaces with the pitch of pure tones. However, the ability to crossmodally match stimuli is presumably possible between any pair of prothetic stimulus dimensions, only a few of which might be argued to pick-up on the same stimulus, or perceptual, property (cf. Cohen, 1934; Mellers & Birnbaum, 1982).

¹⁷ On sight-sound translation, von Hornbostel observed: "Strange! It is just where eye and ear differ, that their connection has been sought. For colour, like pitch, changes with the wavelength: a fact which misled the physicists. (In reality there is a difference at this very point: brightness in hearing depends on frequency, while brightness in sight depends on amplitude.)" (Von Hornbostel, 1938, p. 212).

¹⁸ Mellers and Birnbaum's (1982) talk of 'different continua' presumably supports the view of modality-specific (though comparable) perceptual continua rather than supporting an amodal account.

¹⁹ The notion of an absolute correspondence, implying a unique and direct mapping between two stimuli presented in different sensory modalities can be contrasted with relative correspondences, where the crossmodal mapping is not fixed but depends on the comparison set.

To summarize, beyond shedding light on the way in which stimuli are organized within distinct perceptual dimensions in discrete sensory modalities, the distinction between meta-thetic and prothetic stimulus dimensions leads to the related distinction between absolute versus relative crossmodal correspondences. The latter distinction is, in turn, instrumental when it comes to assuming the existence of amodal concepts, conceived of as the same physical property (such as shape) being picked up via multiple senses (see Lewkowicz & Turkewitz, 1980). This, it should be noted, is subtly different from von Hornbostel's (1931) notion of universal dimensions of perceptual experience. The emphasis in the latter case would appear to be on the perceptual experience itself (i.e., what it is like), whereas the emphasis for many of the amodal dimensions that have been proposed has been on the multiple routes to picking-up information about physical properties out there, regardless of the perceptual qualities that may be associated with that information.

In addition to these early objections, a number of additional issues have also been raised in the literature regarding the status of amodal properties (see Spence & Di Stefano, *submitted*). On the one hand, there would appear to be disagreement about how, exactly, amodal dimensions should be defined. As observed by Johnstone (2021), even Aristotle left it somewhat unclear as to whether these common sensibles should be common to all of the senses or just to two or more of them (see also Bahrack, 2009; Gogate & Bahrack, 1998; Walker-Andrews, 1994). Aligning with most commentators (e.g., Knuuttila, 2008), Johnstone takes Aristotle's considered view to have been that common sensibles are perceptible in their own right by more than one sensory modality, but need not necessarily be perceptible by all five of the commonly accepted senses. Some commentators have suggested that the same dimension, such as sensory intensity (Lewkowicz & Turkewitz, 1981),²⁰ sensory brightness (von Hornbostel, 1931; though see Cohen, 1934), or even sensory 'thickness' (Moul, 1930) should be considered as amodal dimensions, given that these perceptual attributes/dimensions can be associated with two or more (and possibly all) of the senses. Others, though, have wondered whether the robust psychophysics (e.g., of transitivity) that is obtained when

comparing judgments across various pairs of senses (Ellermeier et al., 2021, on the ratio-based crossmodal matching of visual brightness and sound intensity; cf. Heller, 2021; Luce et al., 2010) might not merely reflect the application of ratio properties within qualitatively distinct unimodal prothetic dimensions (see Cohen, 1934; Root & Ross, 1965; Stevens, 1957, 1966, 1971; Stevens & Guirao, 1963). Of course, if the latter suggestion were to be correct then it should not matter which particular pair of prothetic dimensions are chosen for crossmodal matching/scaling. However, those who support the existence/preferential status of certain specific amodal dimensions of experience would presumably have to predict that crossmodal mapping based on a putatively amodal dimension should be more robust, and perhaps also develop earlier, than those correspondences between two distinct prothetic dimensions.

Another possible approach to amodal dimensions is based on the notion of redundant information. As Gogate and Bahrick (1998) put it: "Amodal information is information which is completely redundant across two or more senses" (p. 99). Here, though, it is important to note that there is virtually never perfect redundancy between the senses, even when multiple senses are potentially capable of picking-up on the same environmental property, such as, for example, size/shape (Spence et al., 2013), the precision/accuracy of different unisensory estimates rarely aligns perfectly (Ernst & Banks, 2002). At the same time, and as has already been mentioned, vision and touch only pick up on the same shape/size information over a very narrow range of stimulus sizes.

Ultimately, it is obviously going to be much easier to translate between the senses if there are amodal (and/or absolute) correspondences rather than if crossmodal matches (correspondences) are relative and/or context-dependent. However, while there is evidence for structural translation (e.g., of temporal patterns) across the senses, other dimensions that have been proposed as amodal are, in fact, based on relative (or relational) judgments instead. Returning to the questions that were raised at the start of this article, it can be seen how the existence of amodal stimulus dimensions should allow for a literal, rather than merely metaphorical, translation of a given property. However, while the temporal structure might allow for the matching of a tactile or visual rhythm with an auditory temporal pattern (and so offers the means of conveying the temporal structure) what cannot so easily be captured is the beat attached to auditory temporal patterns (e.g., Grahm, 2012).

Affective (emotionally mediated) correspondences

Reviewing the literature, it is striking how many of those artists who have attempted to search for perceptually meaningful correspondences between colour and music have

²⁰ So, what dimensions count as amodal? According to Lewkowicz and Turkewitz (1980): "Intensity, rate, duration, spatial location, spatial extent, rhythm, and shape all represent amodal features of the world that can be specified in more than one modality. They stand in distinction to such modality specific features of stimulation as redness, sweetness, and pitch" (p. 597). Meanwhile, according to Walker-Andrews (1994): "Amodal properties (size, texture, flexibility, duration, and intensity) that may be picked up by any modality" (p. 42). Von Hornbostel (1931) writes of intensity as a universal dimension of sensory experience (see also Hayek, 1952, p. 21).

ended up stressing the emotional, or affective, basis of the crossmodal matches that they have intuited, or managed to document empirically (e.g., Bragdon, 1916, 1918, p. 139; Cutietta & Haggerty, 1987; Zilczer, 1987; see Hartshorne, 1934; Marin et al., 2012; Spence, 2020a, and Spence & Di Stefano, 2022b, for reviews). In fact, emotional mediation has recently been presented as one of the key factors accounting for a very wide range of different crossmodal correspondences (see Spence, 2020a, for a review).

Support for emotional mediation in the case of audiovisual crossmodal correspondences comes from the results of a study published by Palmer et al. (2013), in which the participants had to associate musical excerpts to colour patches and rate both for their emotional valence (e.g., happy, sad, angry, calm, strong, weak, lively, and dreary). The results highlighted significant correlations between the emotional character of the musical excerpts and those of the colour patches that were chosen to match them (see also Bresin, 2005; Simpson et al., 1956; Whiteford et al., 2018).²¹ Along similar lines, Isbilen and Krumhansl (2016) tested their participants with music excerpts from Bach's *Well-tempered Clavier* and a sample of saturated colours. Their results confirmed that music–color associations can be accounted for by the correlations between music and emotion, and color and emotion. Interestingly, the experimental sample included synaesthetes and those with absolute pitch, who failed to show any peculiar behaviour despite their unusual/extraordinary abilities.

The evidence suggests that emotional mediation is also relevant for those crossmodal associations involving olfaction (e.g., see Levitan et al., 2015, for odour–music associations; Schifferstein & Tanudjaja, 2004; and Gilbert et al., 2016, for colour–taste/smell associations; Di Stefano et al., 2022a). Winter's (2016b) findings might also be taken to indirectly support such a central role of emotion, explaining why it is that taste (gustation), in particular, is a common source domain in most of the crossmodal correspondences, as taste and smell are simply more strongly emotionally valenced (see Levinson & Majid, 2014; Winter, 2016a).

Based on Spence's (2020a) review, it would seem that emotion offers a crucial explanatory concept underpinning the majority of audiovisual associations (see Hartshorne, 1934, for a similar position). That being said, several relevant questions arise here. First, one reviewer of this manuscript wondered why colour should be mapped to emotion in the first place (D'Andrade & Egan, 1974; Jonauskaitė et al., 2020). In this case, the answer may well relate to the fact that exposure to different colours has been documented to

directly affect people's emotions (Jonauskaitė et al., 2019). So, for example, exposure to a Baker-Miller Pink environment may help to calm people (Schauss, 1981, 1985; though see Genschow et al., 2015). Relatedly, it is legitimate to ask whether emotional associations can also be invoked to explain mappings between pitch, brightness, and loudness. This would seem unlikely, but returning to the debate between von Hornbostel, Cohen, and Hartshorne discussed earlier, this might better be considered as a relational (or relative) correspondence, i.e., a kind of analogical mapping (Ravignani & Sonnweber, 2017). Readers wanting to know more about the importance of emotional mediation to explaining many audiovisual crossmodal correspondences are directed to the reviews by Spence and Di Stefano (2022a, 2022b), where the emotional mediation account of audiovisual correspondences is discussed in much more detail.

Problems for any attempt to translate directly, or 'literally', between the senses

There are a number of further potential problems for anyone wanting to translate between the senses (in particular, between individual sensory impressions); these relate to attempts to extend from the crossmodal pairing of individual auditory and visual stimuli to matches between more complex combinations of sensory stimuli (Marks & Bornstein, 1987).

Intramodal versus crossmodal grouping

Intramodal perceptual grouping (defined in terms of the Gestalt grouping principles; see Wagemans, 2015) tends to be much stronger than crossmodal perceptual grouping (see Spence, 2015, for a review). As such, any association, or correspondence that might be established between a specific colour and a particular sound (Sebba, 1991) is likely to be overridden by the emerging intramodal perceptual grouping that will likely take precedence as soon as more than one stimulus is presented in either modality (cf. Bhattacharya & Lindsen, 2016; Collier & Hubbard, 2001; Cuddy, 1985; Cutietta & Haggerty, 1987; Galeyev, 1976; Platt et al., 1990; Woods et al., 2016; Woods & Spence, 2016). Note that this observation can be seen as following naturally out of the fact that perceptual similarity can be more easily understood when occurring within, but not between the senses, given that 'grouping by similarity' is one of the central Gestalt grouping principles. Similarly, consider here only how a given sequence of musical notes may be associated with positive emotion if the sequence is arranged as an ascending pitch series, but the same sounds when organized as a descending sequence, is associated with negative emotion instead. In all these cases,

²¹ Bear in mind here also that people often associate emotions with individual sensory stimuli, such as colours, tastes, and shapes (Wang & Zhang, 2015; Zhou & Tse, 2020).

the meaning is linked to structure of elements or emergent property (see also Cuddy, 1985).

As soon as one starts to look at the correspondence between works of art (e.g., when looking for correspondences between pieces of music and paintings), then the influence of cross-media artistic styles starts to become increasingly relevant (e.g., Arnheim, 1974, 1986; Dailey et al., 1997). In such cases, the audiovisual correspondence may be based on the higher-level structural processes due to language, culture, abstract symbolization, learning, rather than any particular association that may exist between the individual component stimuli (e.g., colours/shapes or musical notes; e.g., Actis-Grasso et al., 2017; Adams, 1995; Albertazzi et al., 2015, 2020; Duthie, 2013; Duthie & Duthie, 2015). Some have even referred to aesthetic correspondences between the arts (Schueller, 1953).²² Notice how, in all such cases, it's the Gestalt organization (or artistic style) that likely dictates the crossmodal matches that are deemed most appropriate. According to O'Malley (1957):

Interrelation of the arts, taken somewhat for granted in most discussions of aesthetic movements or tempers, tend to resist exact definition because they depend more on vague complex similarities in the general aims and expressive ideals of artistic generations than on easily recognizable resemblances among elements of the several arts. Even in attempts to compare such elements (as in equating figurative design with melody, or coloring with harmony), the analogy refers essentially to broad similarities in formal functions, not to specific resemblances between impressions of different senses. Care should be taken, therefore, to distinguish between the correspondence of the arts and intersense analogy. (p. 402).

Similar observations also hold for the combination of soundtracks with movies, with studies showing that composer-intended music-film combinations tend to be selected by participants as providing the best fit (Lipscomb & Kendall, 1994). However, in such cases, it may be more of an affective match (cf. de Staël, 1869, pp. 485–486; O'Malley, 1957, p. 403). Furthermore, the apparent synchronization of the component stimuli may also play an important role as

²² The perceived structure of abstract paintings has been shown to be influenced by the structure of music listened to on initial viewing (Lindner & Hynan, 1987). By perceived structure, Lindner and Hynan (1987) have in mind whether the music is more or less organized (they contrast highly unstructured avant-garde music vs. highly structured minimalistic music, in particular the first 8 minutes of Bulent Arel's "Stereo Electronic Music #1" vs. Philip Glass's "Contrary Motion"). Those participants who heard the minimalistic music later rated the paintings as controlled, ordered, and unified, whereas those who hears the avant-garde music rated the paintings as more accidental, chaotic, and fragmented (cf. Parrott, 1982; Wehner, 1966).

well (see Daurer, 2010; Muller, 2010). At the same time, however, Stechow (1953, p. 324) notes how "vague associations between music and architecture are not very rare. In a sense, Moussorgsky's Gate of Kiev can be considered as such rather than as a translation into music of Victor Hartmann's drawing (Frankenstein, 1939)." Any correspondence that is experienced in the latter case might well be considered to be based of transfer from the temporal aspects of (auditory) to the spatial (visual) domain (Julesz & Hirsch, 1972).

Direct association versus perceptual inference

Intriguing work from Schloss et al. (2018) has highlighted the fact that people's interpretation of the 'meaning' of a given colour may sometimes be determined not by the strength of any direct crossmodal mapping but rather by whatever other stimuli happen to be in the comparison set. As such, colour mappings (and presumably any other kind of crossmodal mapping) may sometimes be inferred rather than necessarily signalling the strongest possible association between the component stimuli that are available for comparison (Mukherjee et al., 2022). As such, in judging the efficacy of any attempt to translate between the senses, it may be important for the interpreter (that is, the person trying to make sense of the stimuli) to know the intention of whoever came up with the translation scheme (i.e., the code mapping the one sensory stimulus to another, and/or the range of stimuli) in the first place, and also for there to assume a communicative, or signalling, function behind the selection, or choice, of colours. This notion is referred to as 'semantic discriminability theory' by Mukherjee and colleagues. There may also be a link to the literature on colour-in-context theory here (Elliott & Meier, 2012, 2014). According to the latter account, the meaning that is attached to a given colour depends on the context in which it is presented. Think only of how red primes temperature (i.e., hot) in the context of taps (where the contrast is with blue) whereas red primes stop/danger when in the context of traffic (and when the comparison may be green; e.g., in traffic lights). So, even if the comparison set does not provide a context as such (e.g., in the sense of 'colour-in-context' theory), nevertheless, at least according to semantic discriminability theory, participants' choices might be constrained by the range of stimuli people are given to respond with. Once again, note how these concerns argue against an absolute mapping between the senses, and thus an easy translation of one sense to another.

Transduction, mimesis, analogy, and parallelism

We conclude by briefly mentioning a few other notions that have not been touched on in this review, but might be considered relevant when attempting to explain sensory

translation. First, transduction (Helmreich, 2015a, b; Lick, 2022; Newfield, 2017), which can be defined as the process taking place when many sensors in the body convert physical signals from the environment into encoded neural signals sent to the central nervous system (Schacter et al., 2010). It should, though, be noted that transduction, strictly defined, is a biologically-determined process as there is not really any choice about the conversion of sensory input in a particular pattern of neural signals (though see also Culache, 2015a, 2015b). Taking a broader perspective, Lick (2022, p. 6) suggests that: “Transductions have been studied within different fields and research areas, for example translation studies, built environment, and education. In general, when adopting a multimodal discursive approach, the term “transduction” (Kress, 2010), also referred to as “resemiotization” (Iedema, 2003), pertains to situations where meaning is shifted from one mode to another, such as written information in a report, which is visualized in a diagram (writing vs. picture). Whereas this transduction uses modes from the same modality (visual), transductions may also be performed by changing from one modality to another, like a presenter’s speech which is simultaneously projected on the wall behind them (auditory vs. visual; Jewitt et al., 2016; Kress, 2010). It must be mentioned that in any transduction the overall meaning of the multimodal text needs to be maintained to ensure the intended interpretation processes (Culache, 2015a, b).”

The notion of mimesis has been evoked by Connor (2004) to account for the relations between sound and touch as follows: “The relations between sound and touch ... tend to be mimetic: Touch accompanies, mimics, performs sound rather than translating ... it” (p. 154; cf. Taussig, 1993, on the notion of mimesis). By contrast, sensory translation occurs primarily, and more properly, between sound and sight, with the information provided by sight being the transformation of the one obtained auditorily. It should also be stressed that Piesse (1867, 1891), the chemist and perfumer, would appear to have been more interested in drawing analogies between the ways in which elements within the auditory and olfactory modalities could be combined harmoniously (see also Cooke & Myin, 2011), rather than necessarily on establishing any direct crossmodal perceptual match (or assert any form of perceptual similarity) between individual auditory and olfactory stimuli. Several commentators have thus referred to the notion of colour-tone analogies (Gombrich, 1979; Jewanski, 2010b; Jewanski & Naumann, 2010; see also O’Malley, 1957).²³ Some years ago

now, Stechow (1953) had already highlighted an important distinction between different kinds of relation: “*translations* from the visual arts into music and *parallelisms* between the visual arts and music” (p. 324, italics in original). Later, he observed that “it would seem to me that comparability of structure reveals a more ‘real’ relationship between such works of art than a mere affinity of ‘mood’ or ‘texture’ could suggest” (Stechow, 1953, p. 325). The latter comment presumably emphasizing the structural rather than the affective nature of correspondences.

Conclusions

As this narrative historical review has hopefully made clear, people have been interested in translating between the senses for millennia (see also Spence & Di Stefano, 2022a, b, for reviews). Beyond philosophers and artists, researchers have also been attracted by sensory translation in their attempt to explain its underlying psychological mechanisms. Several perceptual phenomena have been evoked, such as synaesthesia and crossmodal correspondences. And while traditional synaesthetic and structural mapping approaches have largely failed to explain sensory translation (at least they have failed in the sense of not providing a broadly consensual crossmodal mapping; see Jewanski, 2010a), crossmodal correspondences would appear to offer an alternative way of thinking about the translation between the senses (see Arnheim, 1986, for a broadly similar conclusion).

However, the majority of the evidence that has been published to date suggests that an approach based on affective, or emotionally-mediated, crossmodal mappings is more likely to work than attempts to search for perceptual (possibly amodal) correspondences that are based on putative crossmodal perceptual similarity instead (cf. Hartshorne, 1934; Spence, 2020a; von Helmholtz, 1878/1971; though see also Marks, 2011). Ultimately, therefore, it can be argued that the best that one can hope for as far as matching, or translating between, the senses is to use emotionally mediated correspondences (e.g., Cunningham & Weinel, 2016; Hauck et al., 2022; Spence, 2020a; Spence & Di Stefano, 2022a, for a review). Evidence showing that basic emotions are recognized in musical stimuli across cultures (Fritz et al., 2009) would at least partially support the alleged universality of some elements of emotion-mediated translation from audition and vision. That being said, though, there are presumably only a relatively limited range of distinct emotions (or emotional states) to play with, thus likely limiting the range of possible translations.

Returning, then, to the three key research questions that were outlined in the Introduction: (1) How the topic of sensory translation is related to synaesthesia, multisensory integration and crossmodal associations? (2) Are there common

²³ O’Malley (1957) ponders: “But how could the spectrum-octave parallel ever have been considered significant for the supposed perceptibility of color–sound analogies? To take it that way was to confuse the physical and psychological, as Rousseau and others in the eighteenth century pointed out” (p. 401).

processing mechanisms across the senses that guarantee for sensory translation or, rather, is mapping among the senses based on allegedly universal stimulus dimensions (e.g., amodal)? (3) Is the term ‘translation’ in the context of sensory mapping used metaphorically or literally? In answer to the first question, the topic of sensory translation is related to the topic of synaesthesia because the vivid concurrents experienced by synaesthetes have often been considered as providing guidelines for appropriate translation, especially in the field of arts (seemingly neglecting the fact that synaesthesia is defined by the idiosyncrasy of the inducer-concurrent relations). Meanwhile, the link between sensory translation, multisensory integration, and crossmodal associations can be seen in terms of the emerging literature demonstrating that crossmodal correspondences both modulate multisensory integration and also provide a more consensual guide to translating between the senses (see Pinardi et al., 2023). Regarding the second question, the reviewed literature seems insufficient to support the existence of processing mechanisms that guarantee for sensory translation across the senses. Rather, it would seem to suggest that the only common processing mechanisms across the senses may relate to magnitude (e.g., Pinel et al. 2004; Walsh, 2003; though see also Ronga et al., 2012), though further, and more solid, evidence is required in this direction. At the same time, although universal stimulus dimensions (e.g., amodal or intersensory) have been proposed by a number of researchers over the years, convincing empirical support for their existence has not been forthcoming. Finally, the answers to the first two questions provide indications to address the third one. Observing that, with the possible exception of some theoretically extreme positions (e.g., Paradis & Eeg-Okofsson, 2013; Rakova, 2003), the term ‘translation’ when used in the context of sensory mapping would appear to be used metaphorically rather than literally, thus indicates that the relationship between the sensory impressions that are being ‘translated’ is based on a merely perceptual basis, being it semantic, emotional or cognitive.

While the discussion in the latter parts of this review has focused specifically on attempts to translate between auditory and visual stimuli, it is worth noting that a growing number of crossmodal correspondences have now been documented between the other senses as well (e.g., olfaction, gustation, and touch; e.g., Belkin et al., 1997; Crisinel & Spence, 2012b; Di Stefano et al., 2022a; Gilbert et al., 1996,

2016; Kemp & Gilbert, 1997; Motoki et al., 2020; Motoki et al., 2022; Piesse, 1867, 1891; Raevskiy et al., 2022; Spence, 2020c; Spence et al., 2015; Watson & Gunter, 2017). However, here again, in those cases where a specific source object cannot be identified,²⁴ emotional mediation appears to provide the most parsimonious explanation for the various mappings (e.g., between hue and olfaction, or between hue and colour) that have been documented to date (see Gilbert et al., 2016; Schifferstein & Tanudjaja, 2004).

Directions for future research

People appear to show broad agreement regarding matching in the cases of the crossmodal correspondences, such as those correspondences that have been documented between auditory and visual stimuli. It has also been suggested that they also have a ‘feeling of knowing’ what the consensual (i.e., consensual in the sense of commonly agreed, given that there is no objectively correct answer) answer is likely to be (Koriat, 1975, 1976, 2008, 2011; Rader & Tellegen, 1987).²⁵ In the future, it will be interesting to further study consensuality across a given population as well as its consistency within an individual over time, as these may both be considered to provide measures of the strength of such correspondences. Further research is also needed in order to explain why it is that certain correspondences appear to be stronger, or more robust, than others (cf. Parise, 2016). While as yet there has been little research on such issues, it appears that crossmodal correspondences tend to be fairly consistent (or stable) over time (e.g., Belkin et al., 1997; O’Mahony, 1983). It is interesting to consider what the link might be between the strength/consensuality of crossmodal correspondences and the directionality of synaesthetic metaphor. This is undoubtedly an area where additional research is very much needed.

Looking to the future, protocols investigating sensory substitution might shed light on the way the sensory system can manage the transformation of information from one sensory modality into another (see Pinardi et al., 2023). It will also be interesting to keep an eye on the machine learning literature in order to see whether such big data approaches are capable of turning up any crossmodal matches that work better (in the sense of being more consensual) than those that have been uncovered to date that been based on intuition, synaesthesia, structural alignment, or, on occasion, experimentation (Murari et al., 2020; see also Conway & Christiansen, 2006).

²⁴ For example, when yellow is picked as the corresponding colour for the smell of citrus because both are linked to the same source object, a lemon (see Spence, 2020d).

²⁵ This is distinct from what is seen in the case of synaesthesia.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abboud, S., Hanassy, S., Levy-Tzedek, S., Maidenbaum, S., & Amedi, A. (2014). EyeMusic: Introducing a “visual” colourful experience for the blind using auditory sensory substitution. *Restorative Neurology & Neuroscience*, *32*, 247–257.
- Actis-Grasso, R., Zavagno, D., Lega, C., Zani, A., Daneyko, O., & Cattaneo, Z. (2017). Can music be figurative? Exploring the possibility of crossmodal similarities between music and visual arts. *Psihologija*, *50*(3), 285–306. <https://doi.org/10.2298/PSI1703285A>
- Adams, C. S. (1995). Artistic parallels between Arnold Schoenberg's music and painting (1908–1912) *College Music Symposium*, *35*, 5–21. <https://www.jstor.org/stable/40374266>.
- Adeli, M., Rouat, J., & Molotchnikoff, S. (2014). Audiovisual correspondence between musical timbre and visual shapes. *Frontiers in Human Neuroscience*, *8*, 352. <https://doi.org/10.3389/fnhum.2014.00352>
- Aguiar, D., & Queiroz, J. (2013). Semiosis and intersemiotic translation. *Semiotica*, *196*, 283–292. <https://doi.org/10.1515/sem-2013-0060>
- Ahsen, A. (1997). Visual imagery and performance during multi-sensory experience, synaesthesia and phosphenes. *Journal of Mental Imagery*, *21*(3/4), 1–40.
- Akinbiyi, T., Reiley, C. E., Saha, S., Burschka, D., Hasser, C. J., Yuh, D. D., & Okamura, A. M. (2006). Dynamic augmented reality for sensory substitution in robot-assisted surgical systems. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 567–570). IEEE.
- Albertazzi, L., Canal, L., & Micciolo, R. (2015). Cross-modal association between materic painting and classical Spanish music. *Frontiers in Psychology*, *6*, 424. <https://doi.org/10.3389/fpsyg.2015.00424>
- Albertazzi, L., Canal, L., Micciolo, R., & Hachen, I. (2020). Cross-modal perceptual organization in works of art. *i-Perception*, *11*(4), 1–22. <https://doi.org/10.1177/2041669520950750>
- Alves, B. (2005). Digital harmony of sound and light. *Computer Music Journal*, *29*(4), 45–54. <https://doi.org/10.1162/014892605775179982>
- Anderson, K. (1980). Two translations of Baudelaire's “Correspondances” *Comparative Literature Studies*, *17*(4), 439–446.
- Anikin, A., & Johansson, N. (2019). Implicit associations between individual properties of color and sound. *Attention, Perception, & Psychophysics*, *81*, 764–777. <https://doi.org/10.3758/s13414-018-01639-7>
- Argelander, A. (1927). *Das Farbenhören und der synästhetische der Wahrnehmung* [Coloured hearing and synaesthetic perception], Fischer
- Aristotle. (1907). *De Anima*, transl. by R. D. Hicks. Cambridge University Press.
- Aristotle. (1908). *The Parva Naturalia*, transl. by J. A. Smith, & W. D. Ross. Oxford, UK: Clarendon Press.
- Arnheim, R. (1974). Colors: Irrational and rational. *The Journal of Aesthetics and Art Criticism*, *33*(Winter), 149–154.
- Arnheim, R. (1986). *New essays on the psychology of art*. University of California Press.
- Bahrack, L. E. (2009). Amodal perception. In E. B. Goldstein (Ed.), *Encyclopedia of perception* (pp. 44–46). SAGE Publications.
- Bahrack, L. E. (2010). Intermodal perception and selective attention to intersensory redundancy: Implications for typical social development and autism. *Blackwell Handbook of Infant Development*, *1*, 120–165.
- Bahrack, L. E., & Pickens, J. N. (1994). Amodal relations: The basis for intermodal perception and learning. In D. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp. 205–233). Erlbaum.
- Bahrack, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions in Psychological Science*, *13*(3), 99–102. <https://doi.org/10.1111/j.0963-7214.2004.00283.x>
- Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2012). Language and simulation in conceptual processing. In M. De Vega, A. Glenberg, & A. Graesser (Eds.), *Symbols and embodiment: Debates on meaning and cognition* (pp. 245–283). Oxford University Press.
- Baudelaire, C. (1857). *Les fleurs du mal* [Flowers of evil]. Poulet-Malassis.
- Baudelaire, C. (1954). *Les fleurs du mal* [Flowers of evil] [W. Aggeler, Trans.]. Academy Library Guild.
- Belkin, K., Martin, R., Kemp, S. E., & Gilbert, A. N. (1997). Auditory pitch as a perceptual analogue to odor quality. *Psychological Science*, *8*(4), 340–342. <https://doi.org/10.1111/j.1467-9280.1997.tb00450.x>
- Bernstein, I. H., Eason, T. R., & Schurman, D. L. (1971). Hue-tone interaction: A negative result. *Perceptual and Motor Skills*, *33*(3_ suppl), 1327–1330. <https://doi.org/10.2466/pms.1971.33.3f.1327>
- Bhattacharya, J., & Lindsen, J. P. (2016). Music for a brighter world: Brightness judgment bias by musical emotion. *PLoS One*, *11*(2), e0148959.
- Bolognesi, M., & Strik Lievers, F. (2018). How language and image construct synaesthetic metaphors in print advertising. *Visual Communication*, *10*. <https://doi.org/10.1177/1470357218782001>
- Bragdon, C. (1916). Song and light. *The Architectural Record*, *4*(9), 169–172.
- Bragdon, C. (1918). *Architecture and democracy*. Lector House.
- Bresin, R. (2005). What is the color of that music performance? *Proceedings of the International Computer Music Conference* (pp. 367–370). International Computer Music Association.
- Caballero, R. (2009). Cutting across the senses: Imagery in winespeak and audiovisual promotion. In C. Forceville & E. Urios-Aparisi (Eds.), *Multimodal metaphors* (pp. 73–94). Mouton de Gruyter.
- Caivano, J. L. (1994). Color and sound: Physical and psychophysical relations. *Color Research and Application*, *19*(2), 126–133. <https://doi.org/10.1111/j.1520-6378.1994.tb00072.x>
- Castel, L.-B. (1725). Clavecin pour les yeux [Harpisicord for the eyes]. *Mercure de France* (pp. 2552–2577).
- Castel, L.-B. (1726a). Demonstration geometrique du clavecin pour les yeux et pour tous les sens [Geometric demonstrations of harpsicord for the eyes and for all the senses]. *Mercure de France* (pp. 277–292).
- Castel, L.-B. (1726b). Difficultes sur le clavecin oculaire, avec leurs reponses [Difficulties about the ocular harpsicord, with their responses]. *Mercure de France* (pp. 455–465).
- Castel, L.-B. (1735). Nouvelles experiences d'optique & d'acoustique [Novel experiences of optics and acoustics]. *Mémoires pour*

- l'Historie des Sciences et des Beaux Arts* (pp. 1444–1482). [Cited in Marks (1982)].
- Castel, L.-B. (1740). *L'optique des couleurs, fondée sur les simples observations, et tournée sur-tout à la pratique de la peinture, de la teinture et des autres arts coloristes* [Colour optics, founded on simple observations, and their implications for the practice of painting, colouring, and the other colour arts]. Briasson. [Cited in Marks (1982)].
- Castel, L. B. (1751). Du clavecin oculaire [The ocular harpsicord]. *Mercure de France* (pp. 7–21).
- Cazeaux, C. (2002). Metaphor and the categorization of the senses. *Metaphor and Symbol*, 17(1), 3–26. https://doi.org/10.1207/S15327868MS1701_2
- Cecon, K. (2011). Chemical translations: The case of Robert Boyle's experiments on sensible properties. *Annals of Science*, 68(2), 179–198. <https://doi.org/10.1080/00033790.2010.483016>
- Chan, K. Q., Tong, E. M. W., Tan, D. H., & Koh, A. H. Q. (2013). What do love and jealousy taste like? *Emotion*, 13(6), 1142–1149. <https://doi.org/10.1037/a0033758>
- Chen, Y.-C., & Spence, C. (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology: Human Perception and Performance*, 37(5), 1554–1568. <https://doi.org/10.1037/a0024329>
- Clark, H. H. (1970). The primitive nature of children's relational concepts. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 269–278). Wiley.
- Cohen, N. E. (1934). Equivalence of brightness across modalities. *American Journal of Psychology*, 46, 117–119. <https://doi.org/10.2307/1416240>
- Collier, W. G., & Hubbard, T. L. (2001). Musical scales and evaluations of happiness and awkwardness: Effects of pitch, direction, and scale mode. *American Journal of Psychology*, 114(3), 355–375. <https://doi.org/10.2307/1423686>
- Connor, S. (2004). Edison's teeth: Touching hearing. In V. Erlmann (Ed.), *Hearing cultures: Essays on sound, listening, and modernity* (pp. 153–172). Berg.
- Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities pitting abstract against stimulus-specific representations. *Psychological Science*, 17(10), 905–912. <https://doi.org/10.1111/j.1467-9280.2006.01801.x>
- Cooke, E., & Myin, E. (2011). Is trilled smell possible? How the structure of olfaction determines the phenomenology of smell. *Journal of Consciousness Studies*, 18(11/12), 59–95.
- Crisinel, A.-S., & Spence, C. (2012a). Assessing the appropriateness of 'synaesthetic' messaging on product packaging. *Food Quality and Preference*, 26(1), 45–51. <https://doi.org/10.1016/j.foodqual.2012.03.009>
- Crisinel, A.-S., & Spence, C. (2012b). A fruity note: Crossmodal associations between odors and musical notes. *Chemical Senses*, 37, 151–158. <https://doi.org/10.1093/chemse/bjr085>
- Cuddy, L. L. (1985). The color of melody. *Music Perception*, 2(3), 345–360. <http://www.jstor.org/stable/40285303>.
- Culache, O. (2015). Semiotic study on translating modes within multimodal messages: Developing a comparative analysis model of transduced meanings. In S. Costreie & M. Dumitru (Eds.), *Meaning and truth* (pp. 51–64). Pro Universitaria.
- Culache, O. (2015). Transduction and meaning-making issues within multimodal messages. *Symposion*, 2(4), 495–504. <https://doi.org/10.5840/symposion20152432>
- Cunningham, S., & Weinel, J. (2016). The sound of the smell (and taste) of my shoes too: Mapping the senses using emotion as a medium. In *AM '16: Proceedings of the Audio Mostly 2016* (pp. 28–33). ACM Press. <https://doi.org/10.1145/2986416.2986456>
- Curwen, C. (2018). Music-colour synaesthesia: Concept, context and qualia. *Consciousness and Cognition*, 61, 94–106. <https://doi.org/10.1016/j.concog.2018.04.005>
- Cutieta, R. A., & Haggerty, K. J. (1987). A comparative study of color association with music at various age levels. *Journal of Research in Music Education*, 35(2), 78–91. <https://doi.org/10.2307/3344984>
- Cytowic, R. E. (1989a). *Synaesthesia: A union of the senses*. Springer.
- Cytowic, R. E. (1989b). Synesthesia and the mapping of subjective sensory dimensions. *Neurology*, 39, 849–950. <https://doi.org/10.1212/wnl.39.6.849>
- Cytowic, R. E. (1993). *The man who tasted shapes*. G. P. Putnam's Sons.
- Cytowic, R. E., & Wood, F. B. (1982). Synaesthesia II: Psychophysical relations in the synaesthesia of geometrically shaped taste and colored hearing. *Brain and Cognition*, 1(1), 36–49. [https://doi.org/10.1016/0278-2626\(82\)90005-7](https://doi.org/10.1016/0278-2626(82)90005-7)
- D'Andrade, R., & Egan, M. (1974). The colors of emotions. *American Ethnologist*, 1, 49–63.
- Dailey, A., Martindale, C., & Borkum, J. (1997). Creativity, synesthesia and physiognomic perception. *Creativity Research Journal*, 10, 1–8. https://doi.org/10.1207/s15326934crj1001_1
- Daniels, D., & Naumann, S. (2015). Introduction. In D. Daniels & S. Naumann (Eds.), *See this sound: Audiovisiology An interdisciplinary survey of audiovisual culture* (pp. 5–16). Walter König.
- Daniels, D., Naumann, S., & Thoben, J. (2010). See this sound: Audiovisiology. *An interdisciplinary survey of audiovisual culture*. Walter König.
- Daurer, G. (2010). Audiovisual perception. In D. Daniels, S. Naumann, & J. Thoben (Eds.), *See this sound: Audiovisiology An interdisciplinary survey of audiovisual culture* (pp. 329–337). Walter König.
- Dauriac, L. (1902). Des images suggérées par'audition musicale' [Images suggested by musical sounds]. *Revue Philosophique*, 54(October), 488–503.
- Davis, J. W. (1979). A response to Garner's observations on the relationship between colour and music. *Leonardo*, 12(3), 218. <https://doi.org/10.2307/1574213>
- Day, S. (1996). Synaesthesia and synaesthetic metaphors. *Psyche*, 2(32), 1–16.
- de Parville, H. (1883). Association of colors with sounds. *The Popular Science Monthly* (pp. 490–492). https://en.wikisource.org/wiki/Popular_Science_Monthly/Volume_23/August_1883/Association_of_Colors_with_Sounds
- de Staël, M. (1869). *De L'Allemagne* (with pref. by M. X. Marmier). Firmin-Didot et c.: Paris.
- De Ullmann, S. (1945). Romanticism and synaesthesia: A comparative study of sense transfer in Keats and Byron. *PMLA*, 60(3), 811–827.
- Denham, J. (2017, March 31). This artist with synesthesia sees colors in music and paints your favorite songs. *Vice*. <https://www.vice.com/en/article/gyxq73/melissa-mccracken-synesthesia-painter-interview>
- Deroy, O., & Spence, C. (2013). Why we are not all synesthetes (not even weakly so). *Psychonomic Bulletin & Review*, 20, 643–664. <https://doi.org/10.3758/s13423-013-0387-2>
- Di Stefano, N., & Spence, C. (2022). Roughness perception: A multi-sensory/crossmodal perspective. *Attention, Perception, & Psychophysics*, 84, 2087–2114.
- Di Stefano, N., & Spence, C. (2023). Perceptual similarity: Insights from crossmodal correspondences. *Review of Philosophy and Psychology*. <https://doi.org/10.1007/s13164-023-00692-y>
- Di Stefano, N., Murari, M., & Spence, C. (2022a). Crossmodal correspondences in art and science. Odours, poetry, and music. In N. Di Stefano & M. T. Russo (Eds.), *Olfaction: An interdisciplinary perspective from philosophy to life sciences, human perspectives*

- in health sciences and technology (pp. 155–189). Springer. https://doi.org/10.1007/978-3-030-75205-7_10
- Di Stefano, N., Vuust, P., & Brattico, E. (2022b). Consonance and dissonance perception. A critical review of the historical sources, multidisciplinary findings, and main hypotheses. *Physics of Life Reviews*, 43, 272–304. <https://doi.org/10.1016/j.plrev.2022.10.004>
- Donnell-Kotrozo, C. (1978). Intersensory perception of music: Color me trombone. *Music Educators Journal*, 65(December), 32–37. <https://doi.org/10.2307/3395546>
- Downey, J. E. (1912). Literary synaesthesia. *The Journal of Philosophy*, 9, 490–498.
- Dubois, D. (2007). From psychophysics to semiophysics: Categories as acts of meaning. A case study from olfaction and audition, back to colors. In Plümacher, M., Holz, P. (Eds.), *Speaking of colors and odors* (pp. 167–184). John Benjamins. <https://doi.org/10.1075/ceclr.8.10dub>
- Dunne, C. (2014). Watching these spice bags explode is the most satisfying thing you'll do today. *Fast Design*. <https://www.fastcompany.com/3024922/watching-these-spice-bags-explode-is-the-most-satisfying-thing-youll-do-today>
- Duthie, A. C. (2013). *Do music and art influence one another? Measuring cross-modal similarities in music and art*. Iowa State University.
- Duthie, C., Duthie, B. (2015). Do music and art influence one another? Measuring cross-modal similarities in music and art. *Polymath: An Interdisciplinary Arts and Sciences Journal*, 5(1). <https://bradduthie.github.io/Duthie%26Duthie2015.pdf>
- Ekman, G. (1954). Dimensions of color vision. *Journal of Psychology*, 38, 467–474.
- Ekman, G., Engen, T., Kunnapas, T., & Lindman, R. (1964). A quantitative principle of qualitative similarity. *Journal of Experimental Psychology*, 68, 530–536.
- Ellermeier, W., Kattner, F., & Raum, A. (2021). Cross-modal commutativity of magnitude productions of loudness and brightness. *Attention, Perception, & Psychophysics*, 83, 2955–2967. <https://doi.org/10.3758/s13414-021-02324-y>
- Elliot, A. J., & Maier, M. A. (2012). Color-in-context theory. In P. Devine & A. Plant (Eds.), *Advances in experimental social psychology* (4th ed., pp. 61–125). Elsevier. <https://doi.org/10.1016/B978-0-12-394286-9.00002-0>
- Elliot, A. J., & Maier, M. A. (2014). Color psychology: Effects of perceiving color on psychological functioning in humans. *Annual Review of Psychology*, 65, 95–120. <https://doi.org/10.1146/annurev-psych-010213-115035>
- English, H. B. (1923). Colored hearing. *Science*, 57(1476), 444. <https://doi.org/10.1126/science.57.1476.444>
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429–433. <https://doi.org/10.1038/415429a>
- Evans, R. J. W. (1973). *Rudolf II and his world: A study in intellectual history* (pp. 1576–1612). Clarendon Press.
- Ferrari, R. (2015). Writing narrative style literature reviews. *Medical Writing*, 24(4), 230–235. <https://doi.org/10.1179/2047480615Z.000000000329>
- Field, G. (1835). *Chromatics: Or the analogy, harmony, and philosophy of colours*. David Bogue.
- Fishman, A. (2022). The picture looks like my music sounds: Directional preferences in synesthetic metaphors in the absence of lexical factors. *Language and Cognition*, 14(2), 208–227. <https://doi.org/10.1017/langcog.2022.2>
- Flournoy, T. (1893). *Des phénomènes de synopsie [Phenomena of synopsia]*. Alcan.
- Forceville, C. (2006). Non-verbal and multimodal metaphor in a cognitivist framework agendas for research. In G. Kristiansen, M. Achard, R. Dirven, & F. Ruiz de Mendoza Ibáñez (Eds.), *Cognitive linguistics: Current applications and future perspectives* (pp. 379–402). Mouton de Gruyter.
- Frankenstein, A. (1939). Victor Hartmann and Modeste Musorgsky. *The Musical Quarterly*, 25(3), 268–291.
- Franssen, M. (1991). The ocular harpsichord of Louis-Bertrand Castel: The science and aesthetics of an eighteenth-century cause celebre. *Tractrix*, 3, 15–17.
- Frings, C., & Spence, C. (2010). Crossmodal congruency effects based on stimulus identity. *Brain Research*, 1354, 113–122. <https://doi.org/10.1016/j.brainres.2010.07.058>
- Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A. D., & Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology*, 19(7), 573–576. <https://doi.org/10.1016/j.cub.2009.02.058>
- Furley, P., Goldschmied, N. (2021). Systematic vs. narrative reviews in sport and exercise psychology: Is either approach superior to the other? *Frontiers in Psychology*, 12(685082). <https://doi.org/10.3389/fpsyg.2021.685082>
- Galeyev, B. M. (1976). Music-kinetic art medium: On the work of the group 'Prometei' (SKB), Kazan, U.S.S.R. *Leonardo*, 9, 177–182. <https://doi.org/10.2307/1573549>
- Galeyev, B. M. (2003). Evolution of gravitational synesthesia in music: To color and light. *Leonardo*, 36(2), 129–134. <https://doi.org/10.1162/002409403321554198>
- Galeyev, B. M., & Vanechkina, I. L. (2001). Was Scriabin a synesthete? *Leonardo*, 34(4), 357–361. <https://doi.org/10.1162/00240940152549357>
- Gallace, A., Tan, H. Z., & Spence, C. (2007). Multisensory numerosity judgments for visual and tactile stimuli. *Perception & Psychophysics*, 69(4), 487–501. <https://doi.org/10.3758/BF03193906>
- Garner, W. (1978). The relationship between colour and music. *Leonardo*, 11(3), 225–226. <https://doi.org/10.2307/1574153>
- Genschow, O., Noll, T., Wänke, M., & Gersbach, R. (2015). Does Baker-Miller pink reduce aggression in prison detention cells? A critical empirical examination. *Psychology, Crime & Law*, 21(5), 482–489. <https://doi.org/10.1080/1068316X.2014.989172>
- Giannos, K., Athanasopoulos, G., & Cambouropoulos, E. (2021). Cross-modal associations between harmonic dissonance and visual roughness. *Music & Science*, 4, 20592043211055484. <https://doi.org/10.1177/20592043211055484>
- Gibson, E. J. (1969). *Principles of perceptual learning and development*. Appleton.
- Gil, D., & Shen, Y. (2021). Metaphors: The evolutionary journey from bidirectionality to unidirectionality. *Philosophical Transactions of the Royal Society B*, 376(1824), 20200193. <https://doi.org/10.1098/rstb.2020.0193>
- Gilbert, A. N., Martin, R., & Kemp, S. E. (1996). Cross-modal correspondence between vision and olfaction: The color of smells. *American Journal of Psychology*, 109(3), 335–351. <https://doi.org/10.2307/1423010>
- Gilbert, A. N., Fridlund, A. J., & Lucchina, L. A. (2016). The color of emotion: A metric for implicit color associations. *Food Quality & Preference*, 52, 203–210. <https://doi.org/10.1016/j.foodqual.2016.04.007>
- Ginsberg, L. (1923). A case of synaesthesia. *American Journal of Psychology*, 34, 582–589. <https://doi.org/10.2307/1414059>
- Goethe, J. W. V. (1840). *Theory of colours (Charles Locke Eastlake, Trans.)*. John Murray. Original work published 1810.
- Gogate, L. J., & Bahrack, L. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *Journal of Experimental Child Psychology*, 69(2), 133–149. <https://doi.org/10.1006/jecp.1998.2438>
- Gombrich, E. H. (1960). *Art and illusion*. Phaidon Press.
- Gombrich, E. H. (1979). Epilogue: Some musical analogies. *The sense of order: A study in the psychology of decorative art* (pp. 285–305). Phaidon.

- Goodman, N. (1972). Seven strictures on similarity. In N. Goodman (Ed.), *Problems and projects* (pp. 437–446). Bobs-Merrill.
- Gottlieb, H. (2018). Semiotics and translation. In K. Malmkjær (Ed.), *Routledge handbook of translation studies and linguistics* (pp. 45–63). Routledge.
- Gouk, P. M. (2000). Music in Francis Bacon's natural philosophy. In P. Bozza (Ed.), *Number to sound: The musical way to the scientific revolution* (pp. 135–152). Springer.
- Grahn, J. A. (2012). See what I hear? Beat perception in auditory and visual rhythms. *Experimental Brain Research*, 220(1), 51–61. <https://doi.org/10.1007/s00221-012-3114-8>
- Grossenbacher, P. G., & Lovelace, C. T. (2001). Mechanisms of synesthesia: Cognitive and physiological constraints. *Trends in Cognitive Sciences*, 5(1), 36–41. [https://doi.org/10.1016/S1364-6613\(00\)01571-0](https://doi.org/10.1016/S1364-6613(00)01571-0)
- Gurman, D., McCormick, C. R., & Klein, R. M. (2021). Crossmodal correspondence between auditory timbre and visual shape. *Multisensory Research*, 35(3), 221–241. <https://doi.org/10.1163/22134808-bja10067>
- Hamilton-Fletcher, G., Wright, T. D., & Ward, J. (2016). Cross-modal correspondences enhance performance on a colour-to-sound sensory substitution device. *Multisensory Research*, 29(4/5), 337–363. <https://doi.org/10.1163/22134808-00002519>
- Hamilton-Fletcher, G., Witzel, C., Reby, D., & Ward, J. (2017). Sound properties associated with equiluminant colours. *Multisensory Research*, 30(3/5), 337–362. <https://doi.org/10.1163/22134808-00002567>
- Hankins, T. L. (1994). The ocular harpsichord of Louis-Bertrand Castel; Or, the instrument that wasn't. *Osiris*, 9, 141–156. <http://www.jstor.org/stable/302003>.
- Harrison, J. (2001). *Synaesthesia: The strangest thing*. Oxford University Press.
- Hartshorne, C. (1934). *The philosophy and psychology of sensation*. University of Chicago Press.
- Hauck, P., von Castell, C., & Hecht, H. (2022). Crossmodal correspondence between music and ambient color is mediated by emotion. *Multisensory Research*, 35(5), 407–446. <https://doi.org/10.1163/22134808-bja10077>
- Haverkamp, M. (2014). *Synesthetic design: Handbook for a multisensory approach*. Birkhäuser.
- Hayek, F. A. (1952). *The sensory order*. University of Chicago Press.
- Hector, A. B. (1922). *Correlation of colour characters in rice*. Printed and Pub. for the Imperial Department of Agriculture in India, by Thacker, Spink & Company.
- Heller, J. (2021). Internal references in cross-modal judgments: A global psychophysical perspective. *Psychological Review*, 128(3), 509–524. <https://doi.org/10.1037/rev0000280>
- Helmreich, S. (2015a). Transducing. In C. A. Jones, D. Mather, & R. Uchill (Eds.), *Experience: Culture cognition and the common sense* (pp. 163–167). MIT Press.
- Helmreich, S. (2015b). Transduction. In D. Novak & M. Sakakeeny (Eds.), *Keywords in sound* (pp. 222–231). Duke University Press.
- Houston, S. D., & Taube, K. A. (2000). An archaeology of the senses: Perception and cultural expression in ancient Mesoamerica. *Cambridge Archaeological Journal*, 10(2), 261–294. <https://doi.org/10.1017/S095977430000010X>
- Hsiao, J. Y., Chen, Y.-C., Spence, C., & Yeh, S. L. (2012). Assessing the effects of audiovisual semantic congruency on the perception of a bistable figure. *Consciousness and Cognition*, 21, 775–787. <https://doi.org/10.1016/j.concog.2012.02.001>
- Huang, J., Gamble, D., Sarnlertsophon, K., Wang, X., & Hsiao, S. (2012). Feeling music: Integration of auditory and tactile inputs in musical meter perception. *PLOS ONE*, 7(10), e48496. <https://doi.org/10.1371/journal.pone.0048496>
- Huidobro Moya, J. M. (2007). El arte "mágico" de la alquimia [The magic art of alchemy]. *Manual Formativo de ACTA*, 45, 99–106.
- Hunt, H. (2005). Synaesthesia, metaphor and consciousness. *Journal of Consciousness Studies*, 12(12), 6–45.
- Hutmacher, F. (2019). Why is there so much more research on vision than on any other sensory modality?. *Frontiers in Psychology*, 10, 2246.
- Huxley, A. (1932). *Brave new world*. Harper & Row.
- Huysmans, J.-K. (1884/1926). *À rebours*. Charpentier.
- Iedema, R. (2003). Multimodality, resemitotization: Extending the analysis of discourse as multi-semiotic practice. *Visual Communication*, 2(1), 29–57. <https://doi.org/10.1177/1470357203002001751>
- Ione, A., & Tyler, C. (2003). Was Kandinsky a synesthete? *Journal of the History of the Neurosciences*, 12, 223–226. <https://doi.org/10.1076/jhin.12.2.223.15540>
- Ione, A., & Tyler, C. (2004). Neuroscience, history and the arts. Synesthesia: Is F-sharp colored violet? *Journal of the History of the Neurosciences*, 13(1), 58–65. <https://doi.org/10.1080/09647040490885493>
- Isbilen, E. S., & Krumhansl, C. L. (2016). The color of music: Emotion-mediated associations to Bach's Well-tempered Clavier. *Psychomusicology: Music, Mind, and Brain*, 26(2), 149–161. <https://doi.org/10.1037/pmu0000147>
- Itoh, K., Sakata, H., Kwee, I. L., & Nakada, T. (2017). Musical pitch classes have rainbow hues in pitch class-color synesthesia. *Scientific Reports*, 7, 17781. <https://doi.org/10.1038/s41598-017-18150-y>
- Jewanski, J. (2010a). Color organs: From the clavecin oculaire to autonomous light kinetics. In D. Daniels, S. Naumann, & J. Thoben (Eds.), *See this sound: Audiovisiology An interdisciplinary survey of audiovisual culture* (pp. 77–88). Walter König.
- Jewanski, J. (2010b). Color-tone analogies: A systematic presentation of the principles of correspondence. In D. Daniels, S. Naumann, & J. Thoben (Eds.), *See this sound: Audiovisiology An interdisciplinary survey of audiovisual culture* (pp. 338–347). Walter König.
- Jewanski, J., & Naumann, S. (2010). Structural analogies between music and the visual arts. In D. Daniels, S. Naumann, & J. Thoben (Eds.), *See this sound: Audiovisiology An interdisciplinary survey of audiovisual culture* (pp. 388–396). Walter König.
- Jewanski, J., Day, S. A., & Ward, J. (2009). A colourful albino: The first documented case of synaesthesia, by Georg Tobias Ludwig Sachs in 1812. *Journal of the History of the Neurosciences*, 18, 293–303. <https://doi.org/10.1080/09647040802431946>
- Jewanski, J., Simner, J., Day, S. A., & Ward, J. (2011). The development of a scientific understanding of synaesthesia from early case studies (1849–1873). *Journal of the History of the Neurosciences*, 20, 284–305. <https://doi.org/10.1080/0964704X.2010.528240>
- Jewanski, J., Simner, J., Day, S., Rothen, N., & Ward, J. (2020). The "golden age" of synesthesia inquiry in the late nineteenth century (1876–1895). *Journal of the History of the Neurosciences*, 29(2), 175–202. <https://doi.org/10.1080/0964704X.2019.1636348>
- Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge.
- Johnson, H. (2005). *Wine: A life uncorked*. Weidenfeld & Nicolson.
- Johnson, S. (1751, Feb 9). *The Rambler*, 94.
- Johnstone, M. A. (2021). Aristotle on the unity of touch. *Journal of the History of Philosophy*, 59(1), 23–43. <https://doi.org/10.1353/hph.2021.0001>
- Jonauskaitė, D., Abdel-Khalek, A. M., Abu-Akel, A., Al-Rasheed, A. S., Antonietti, J.-P., Åsgeirsson, Á. G., et al. (2019). The sun is no fun without rain: Physical environments affect how we feel about yellow across 55 countries. *Journal of Environmental Psychology*, 66, 101350. <https://doi.org/10.1016/j.jenvp.2019.101350>
- Jonauskaitė, D., Abu-Akel, A., Dael, N., Oberfeld, D., Abdel-Khalek, A. M., Al-Rasheed, A. S., et al. (2020). Universal patterns in

- color-emotion associations are further shaped by linguistic and geographic proximity. *Psychological Science*, 31(10), 1245–1260. <https://doi.org/10.1177/0956797620948810>
- Juhász, A. (1926). A szagerzetek egy új tujolj donsaga. *Athanaeum*, 12, 34–39.
- Julesz, B., & Hirsh, I. J. (1972). Visual and auditory perception - An essay of comparison. In E. E. David & P. B. Denes (Eds.), *Human communication: A unified view* (pp. 283–340). McGraw-Hill.
- Just, D. K. (2017). Was Kandinsky a synaesthete? Examining his writings and other evidence. *Multisensory Research*, 30(3–5), 447–460.
- Kandinsky, W. (1977). *Concerning the spiritual in art, especially in painting* (M. T. H. Sadler, Trans.). Dover Publications.
- Karwoski, T. F., Odbert, H. S., & Osgood, C. E. (1942). Studies in synesthetic thinking. II. The rôle of form in visual responses to music. *Journal of General Psychology*, 26, 199–222.
- Kemp, S. E., & Gilbert, A. N. (1997). Odor intensity and color lightness are correlated sensory dimensions. *American Journal of Psychology*, 110(1), 35–46. <https://doi.org/10.2307/1423699>
- Kiefer, B. (2017). How Guinness and R/GA made a VR tasting experience for all five senses Campaign. <https://www.campaignlive.co.uk/article/guinness-r-ga-made-vr-tasting-experience-five-senses/1434731>
- Kircher, A. (1650). *Musurgia universalis*. Rome.
- Klein Cornwall-Clyne, A. B. (1937). *Coloured light: An art medium* (3rd enlarged ed. of *The art of light: Colour music*). Crosby Lockwood & Son.
- Klutky, N. (1990). Geschlechtsunterschiede in der Gedächtnisleistung für Gerüche, Tonfolgen und Farben [Sex differences in memory performance for odors, tone sequences and colors]. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 37(3), 437–446.
- Knuuttila, S. (2008). Aristotle's theory of perception and medieval Aristotelianism. In S. Knuuttila, & P. Karkkainen (Eds.), *Theories of perception in medieval and early modern philosophy*. Springer Science & Business Media.
- Koriat, A. (1975). Phonetic symbolism and feeling of knowing. *Memory & Cognition*, 3(5), 545–548. <https://doi.org/10.3758/BF03197529>
- Koriat, A. (1976). Another look at the relationship between phonetic symbolism and the feeling of knowing. *Memory & Cognition*, 4(3), 244–248. <https://doi.org/10.3758/BF03213170>
- Koriat, A. (2008). Subjective confidence in one's answers: The consensus principle. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(4), 945–959. <https://doi.org/10.1037/0278-7393.34.4.945>
- Koriat, A. (2011). Subjective confidence in perceptual judgments: A test of the self-consistency model. *Journal of Experimental Psychology. General*, 140(1), 117–139. <https://doi.org/10.1037/a0022171>
- Krantz, D. H. (1972). A theory of magnitude estimation and cross-modality matching. *Journal of Mathematical Psychology*, 9(2), 168–199. [https://doi.org/10.1016/0022-2496\(72\)90025-9](https://doi.org/10.1016/0022-2496(72)90025-9)
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
- Landau, M. J., Meier, B. P., & Keefer, L. A. (2010). A metaphor-enriched social cognition. *Psychological Bulletin*, 136(6), 1045–1067. <https://doi.org/10.1037/a0020970>
- Lavignac, A. (1899). *Music and musicians*. Henry Holt and Company.
- Lee's, C.H. (2013). *Essence in space*, project visible at: <https://www.changheele.com/essence-in-space.html>
- Leibniz, G. W. V. (1896). *On solidity. New essays concerning human understanding*. Macmillan.
- Levinson, S. C., & Majid, A. (2014). Differential ineffability and the senses. *Mind & Language*, 29(4), 407–427. <https://doi.org/10.1111/mila.12057>
- Levitan, C. A., Charney, S. A., Schloss, K. B., & Palmer, S. E. (2015). The smell of jazz: Crossmodal correspondences between music, odor, and emotion. In *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp. 1326–1331).
- Lewkowicz, D. J., & Turkewitz, G. (1980). Cross-modal equivalence in early infancy: Auditory-visual intensity matching. *Developmental Psychology*, 16(6), 597–607. <https://doi.org/10.1037/0012-1649.16.6.597>
- Lewkowicz, D. J., & Turkewitz, G. (1981). Intersensory interaction in newborns: Modification of visual preferences following exposure to sound. *Child Development*, 827–832.
- Liao, M. H. (2018). Museums and creative industries: The contribution of translation studies. *Journal of Specialised Translation*, 29, 45–62.
- Lick, E. (2022). “Multimodal sensory marketing” in retailing: The role of intra- and intermodality transductions. *Consumption Markets & Culture*. <https://doi.org/10.1080/10253866.2022.2046564>
- Lickliter, R., & Bahrick, L. E. (2012). The role of intersensory redundancy in early perceptual, cognitive, and social development. In A. Bremner, D. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 183–206). Oxford University Press.
- Lindner, D., & Hynan, M. T. (1987). Perceived structure of abstract paintings as a function of structure of music listened to on initial viewing. *Bulletin of the Psychonomic Society*, 25(1), 44–46. <https://doi.org/10.3758/BF03330072>
- Lipscomb, S. D., & Kendall, R. A. (1994). Perceptual judgement of the relationship between musical and visual components in film. *Psychomusicology*, 13(1/2), 60–98. <https://doi.org/10.1037/h0094101>
- Locke, J. (1690). *An essay concerning human understanding*. London: Thomas Basset.
- Luce, R. D., Steingrimsson, R., & Narens, L. (2010). Are psychophysical scales of intensities the same or different when stimuli vary on other dimensions? Theory with experiments varying loudness and pitch. *Psychological Review*, 117, 1247–1258. <https://doi.org/10.1037/a0020174>
- Ludwig, V. U., Adachi, I., & Matzuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (*Pan troglodytes*) and humans. *Proceedings of the National Academy of Sciences USA*, 108(51), 20661–20665. <https://doi.org/10.1073/pnas.1112605108>
- MacDougal, R. (1898). Music imagery. A confession of experience. *Psychological Review*, 5(5), 463–476.
- Marin, M. M., Gingras, B., & Bhattacharya, J. (2012). Crossmodal transfer of arousal, but not pleasantness, from the musical to the visual domain. *Emotion*, 12, 618–631. <https://doi.org/10.1037/a0025020>
- Marks, L. E. (1975). On colored-hearing synesthesia: cross-modal translations of sensory dimensions. *Psychological Bulletin*, 82(3), 303.
- Marks, L. E. (1978). *The unity of the senses: Interrelations among the modalities*. Academic Press.
- Marks, L. E. (1982). Bright sneezes and dark coughs, loud sunlight and soft moonlight. *Journal of Experimental Psychology: Human Perception & Performance*, 8(2), 177–193. <https://doi.org/10.1037/0096-1523.8.2.177>
- Marks, L. E. (1982). Synaesthetic perception and poetic metaphor. *Journal of Experimental Psychology: Human Perception and Performance*, 8(1), 15–23. <https://doi.org/10.1037/0096-1523.8.1.15>
- Marks, L. E. (1983). Similarities and differences among the senses. *International Journal of Neuroscience*, 19, 1–12. <https://doi.org/10.3109/00207458309148640>
- Marks, L. E. (1987a). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *Journal of Experimental*

- Psychology: Human Perception and Performance*, 13, 384–394. <https://doi.org/10.1037/0096-1523.13.3.384>
- Marks, L. E. (1987b). On cross-modal similarity: Perceiving temporal patterns by hearing, touch, and vision. *Perception & Psychophysics*, 42(3), 250–256. <https://doi.org/10.3758/BF03203076>
- Marks, L. E. (1996). On perceptual metaphors. *Metaphor and Symbolic Activity*, 11(1), 39–66. https://doi.org/10.1207/s15327868ms1101_3
- Marks, L. E. (2011). Synesthesia, then and now. *Intellectica*, 55, 47–80. <https://doi.org/10.3406/intel.2011.1161>
- Marks, L. E., & Bornstein, M. H. (1987). Sensory similarities: Classes characteristics, and cognitive consequences. In R. E. Haskell (Ed.), *Symbolic structures: The psychology of metaphoric transformation* (pp. 49–65). Ablex.
- Marks, L. E., Hammeal, R. J., & Bornstein, M. H. (1987). Perceiving similarity and comprehending metaphor. *Monographs of the Society for Research in Child Development*, 52(215), 1–102. <https://doi.org/10.2307/1166084>
- Mather, D. (2015). Analogies. In C. A. Jones, D. Mather, & R. Uchill (Eds.), *Experience: Culture cognition and the common sense* (pp. 57–71). MIT Press.
- McAdams, S. (2019). The perceptual representation of timbre. In K. Siedenburg, C. Saitis, S. McAdams, A. Popper, & R. Fay (Eds.), *Timbre: Acoustics, perception, and cognition [Springer Handbook of Auditory Research, Vol 69]*. Springer. https://doi.org/10.1007/978-3-030-14832-4_2
- McAuley, J. D., & Henry, M. J. (2010). Modality effects in rhythm processing: Auditory encoding of visual rhythms is neither obligatory nor automatic. *Attention, Perception, & Psychophysics*, 72(5), 1377–1389. <https://doi.org/10.3758/APP.72.5.1377>
- McKellar, P. (1972). *Imagery and synaesthesia: A sweet taste like middle C*. Science Council of Japan.
- McKellar, P. (1997). Synaesthesia and imagery: *Fantasia* revisited. *Journal of Mental Imagery*, 21(3/4), 41–53.
- Meehan, M., Samuel, L., & Abrahamson, V. (1998). *The future ain't what it used to be: The 40 cultural trends transforming your job, your life, your world*. Riverhead Books.
- Melara, R. D. (1989). Dimensional interaction between color and pitch. *Journal of Experimental Psychology: Human Perception & Performance*, 15(1), 69–79. <https://doi.org/10.1037/0096-1523.15.1.69>
- Mellers, B., & Birnbaum, M. H. (1982). Loci of contextual effects in judgment. *Journal of Experimental Psychology: Human Perception & Performance*, 8(4), 582–601. <https://doi.org/10.1037/0096-1523.8.4.582>
- Menouti, K., Akiva-Kabiri, L., Banissy, M. J., & Stewart, L. (2015). Timbre-colour synaesthesia: Exploring the consistency of associations based on timbre. *Cortex*, 63, 1–3. <https://doi.org/10.1016/j.cortex.2014.08.009>
- Mills, C. B., Boteler, E. H., & Larcombe, G. K. (2003). “Seeing things in my head”: A synesthete’s images for music and notes. *Perception*, 32(11), 1359–1376. <https://doi.org/10.1068/p5100>
- Misdariis, N., Susini, P., Houix, O., Rivas, R., Cerles, C., Lebel, E., Tetienne, A., & Duquesne, A. (2021). *Mapping sound properties and oenological characters by a collaborative sound design approach—Towards an augmented experience*. CMMR—Computer Music Multidisciplinary Research (p. hal-02469362). CMMR Computer Music Multidisciplinary Research.
- Moller, A. C., Elliot, A. J., & Maier, M. A. (2009). Basic hue-meaning associations. *Emotion*, 9, 898–902. <https://doi.org/10.1037/a0017811>
- Moritz, W. (1997, April). The dream of color music, and machines that made it possible. *Animation World Magazine*, 2.1. <https://www.awn.com/mag/issue2.1/articles/moritz2.1.html>
- Motoki, K., & Velasco, C. (2021). Taste-shape correspondences in context. *Food Quality and Preference*, 88, 104082. <https://doi.org/10.1016/j.foodqual.2020.104082>
- Motoki, K., Saito, T., Nouchi, R., & Sugiura, M. (2020). Cross-modal correspondences between temperature and taste attributes. *Frontiers in Psychology*, 11, 571852. <https://doi.org/10.3389/fpsyg.2020.571852>
- Motoki, K., Saito, T., & Velasco, C. (2022). Spontaneous crossmodal correspondences grounded in contexts. *Food Quality and Preference*, 100, 104619. <https://doi.org/10.1016/j.foodqual.2022.104619>
- Moul, E. R. (1930). An experimental study of visual and auditory “thickness” *American Journal of Psychology*, 42(4), 544–560. <https://doi.org/10.2307/1414876>
- Mudge, E. L. (1920). The common synaesthesia of music. *Journal of Applied Psychology*, 4, 342–345. <https://doi.org/10.1037/h0072596>
- Mukherjee, K., Yin, B., Sherman, B. E., Lessard, L., & Schloss, K. B. (2022). Context matters: A theory of semantic discriminability for perceptual encoding systems. *IEEE Transactions on Visualization and Computer Graphics*, 28(1), 697–706. <https://doi.org/10.1109/TVCG.2021.3114780>
- Muller, J. P. (2010). Synchronization as a sound-image relationship. In D. Daniels, S. Naumann, & J. Thoben (Eds.), *See this sound: Audiovisuology An interdisciplinary survey of audiovisual culture* (pp. 401–413). New York: Walter König.
- Müller, N., Nagels, A., & Kauschke, C. (2022). Metaphorical expressions originating from human senses: Psycholinguistic and affective norms for German metaphors for internal state terms (MIST database). *Behavior Research Methods*, 54, 365–377. <https://doi.org/10.3758/s13428-021-01639-w>
- Murari, M., Chmiel, A., Tiepolo, E., Zhang, J. D., Canazza, S., Rodà, A., & Schubert, E., et al. (2020). Key clarity is blue, relaxed, and maluma: Machine learning used to discover cross-modal connections between sensory items and the music they spontaneously evoke. In H. Shoji (Ed.), *KEER 2020, AISC 1256* (pp. 214–223). Springer Nature Singapore Pte.
- Myers, C. S. (1911). A case of synaesthesia. *British Journal of Psychology*, 4, 228–238.
- Myers, C. S. (1914). Two cases of synaesthesia. *British Journal of Psychology*, 7, 112–117.
- Nanay, B. (2018). Multimodal mental imagery. *Cortex*, 105, 125–134. <https://doi.org/10.1016/j.cortex.2017.07.006>
- Nanay, B. (2020). Synesthesia as multimodal mental imagery. *Multisensory Research*, 34(3), 281–296. <https://doi.org/10.1163/22134808-bja10027>
- Neather, R. (2008). Translating tea: On the semiotics of interlingual practice in the Hong Kong museum of tea ware. *Meta: Journal des Traducteurs/Meta: Translators’ Journal*, 53(1), 218–240.
- Neather, R. (2012). Intertextuality, translation, and the semiotics of museum presentation: The case of bilingual texts in Chinese museums. *Semiotica*, 192, 197–218.
- Nelson, M. R., & Hitchon, J. C. (1995). Theory of synesthesia applied to persuasion in print advertising headlines. *Journalism & Mass Communication Quarterly*, 72, 346–360.
- Nelson, M. R., & Hitchon, J. C. (1999). Loud tastes, colored fragrances, and scented sounds: How and when to mix the senses in persuasive communications. *Journalism & Mass Communication Quarterly*, 76, 354–372.
- Newfield, D. (2017). Transformation, transduction and the transmodal moment. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (2nd ed., pp. 100–113). Routledge.
- Newmark, P. (1981). *Approaches to translation*. Pergamon Press.
- Newton, I. (1704). *Opticks*. Available online at: <https://archive.org/details/opticksortreatis00newt>
- Nijis, L., Moens, B., Lesaffre, M., & Leman, M. (2012). The Music Paint Machine: Stimulating self-monitoring through the

- generation of creative visual output using a technology-enhanced learning tool. *Journal of New Music Research*, 41(1), 79–101.
- O'Mahony, M. (1983). Gustatory responses to nongustatory stimuli. *Perception*, 12(5), 627–633. <https://doi.org/10.1068/p120627>
- O'Malley, G. (1957). Literary synesthesia. *The Journal of Aesthetics and Art Criticism*, 15, 391–411.
- O'Regan, J. K. (2011). *Why red doesn't sound like a bell: Understanding the feel of consciousness*. Oxford University Press.
- Ortmann, O. (1933). Theories of synesthesia in the light of a case of colored hearing. *Human Biology*, 5, 155–211.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. University of Illinois Press.
- Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, 15(2), 331.
- Palmer, S. E., Schloss, K. B., Xu, Z., & Prado-León, L. R. (2013). Music–color associations are mediated by emotion. *Proceedings of the National Academy of Sciences*, 110(22), 8836–8841. <https://doi.org/10.1073/pnas.1212562110>
- Paradis, C., & Eeg-Olofsson, M. (2013). Describing sensory experience: The genre of wine reviews. *Metaphor and Symbol*, 28(1), 22–40. <https://doi.org/10.1080/10926488.2013.742838>
- Parise, C. V. (2016). Crossmodal correspondences: Standing issues and experimental guidelines. *Multisensory Research*, 29(1/3), 7–28. <https://doi.org/10.1163/22134808-00002502>
- Parise, C. V., & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: A study using the implicit association test. *Experimental Brain Research*, 220, 319–333. <https://doi.org/10.1007/s00221-012-3140-6>
- Parrott, A. C. (1982). Effects of paintings and music, both alone and in combination, on emotional judgment. *Perceptual and Motor Skills*, 54, 635–641.
- Paterson, M. (2021). *How we became sensorimotor: Movement, measurement, sensation*. University of Minnesota Press.
- Peacock, K. (1985). Synesthetic perception: Alexander Scriabin's color hearing. *Music Perception*, 2, 483–506.
- Peacock, K. (1988). Instruments to perform color-music: Two centuries of technological experimentation. *Leonardo*, 21, 397–406.
- Pedović, I., & Stosić, M. (2018). Předběžná sdělení: A comparison of verbal and sensory presentation methods in measuring crossmodal correspondence within a semantic-based approach. *Československá Psychologie*, 62(6), 602–615.
- Piesse G. W. (1891). *Piesse's art of perfumery* (5th Ed.). Piesse and Lubin. Downloaded from <http://www.gutenberg.org/files/16378/16378-h/16378-h.htm>
- Piesse G. W. S. (1867). *The art of perfumery and the methods of obtaining the odors of plants: With instructions for the manufacture of perfumes for the handkerchief, scented powders, odorous vinegars, dentifrices, pomatums, cosmetics, perfumed soap, etc., to which is added an appendix on preparing artificial fruitescences, etc.* Lindsay & Blakiston.
- Pinel, P., Piazza, M., Le, Bihan, & D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, 41(6), 983–993. [https://doi.org/10.1016/s0896-6273\(04\)00107-2](https://doi.org/10.1016/s0896-6273(04)00107-2)
- Pinardi, M., Di Stefano, N., Di Pino, G., & Spence, C. (2023). Exploring crossmodal correspondences for future research in human movement augmentation. *Frontiers in Psychology*, 14, 1190103.
- Platt, J. R., Racine, R. J., Stark, M., & Weiser, M. (1990). Pitch interactions in the perception of isolated musical triads. *Perception & Psychophysics*, 48(1), 59–67. <https://doi.org/10.3758/bf03205011>
- Plummer, H. C. (1915). Color music—A new art created with the aid of science The color organ used in Scriabine's symphony "Prometheus" *Scientific American*, 112(15), 343. <https://doi.org/10.1038/scientificamerican04101915-343>
- Pomerantz, J. R., Pristach, E. A., & Carson, C. E. (1989). Attention and object perception. In B. E. Shepp & S. Ballesteros (Eds.), *Object perception: Structure and process* (pp. 53–89). Erlbaum.
- Pomp, J., Bestgen, A.-K., Schulze, P., Müller, C. J., Citron, F. M. M., Suchan, B., & Kuchinke, L. (2018). Lexical olfaction recruits olfactory orbitofrontal cortex in metaphorical and literal contexts. *Brain and Language*, 179, 11–21. <https://doi.org/10.1016/j.bandl.2018.02.001>
- Popova, Y. (2005). Image schema and verbal synaesthesia. In B. Hampe (Ed.), *From perception to meaning: Image schemas in cognitive linguistics* (pp. 1–26). Mouton de Gruyter.
- Pridmore, R. W. (1992). Music and color: Relations in the psychophysical perspective. *Color Research & Application*, 17, 57–61. <https://doi.org/10.1002/col.5080170110>
- Quek, Z. F., Schorr, S. B., Nisky, I., Provancher, W. R., & Okamura, A. M. (2015). Sensory substitution and augmentation using 3-degree-of-freedom skin deformation feedback. *IEEE Transactions on Haptics*, 8(2), 209–221. <https://doi.org/10.1109/TOH.2015.2398448>
- Rader, C. M., & Tellegen, A. (1987). An investigation of synesthesia. *Journal of Personality and Social Psychology*, 52, 981–987. <https://doi.org/10.1037/0022-3514.52.5.981>
- Raevskiy, A., Bubnov, I., Chen, Y.-C., & Sakai, N. (2022). Differences in color representations of tastes: Cross-cultural study among Japanese, Russian and Taiwanese. In *International Conference on Human-Computer Interaction* (pp. 378–395). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-06038-0_28
- Rakova, M. (2003). *The extent of the literal: Metaphor, polysemy and theories of concepts*. Palgrave Macmillan.
- Ravignani, A., & Sonnweber, R. (2017). Chimpanzees process structural isomorphisms across sensory modalities. *Cognition*, 161, 74–79. <https://doi.org/10.1016/j.cognition.2017.01.005>
- Reuter, C., Czédik-Eysenberg, I., Siddiq, S., & Oehler, M. (2018a). *The closer the better: The role of formant positions in timbre similarity perception and timbre classification*. In TIMBRE 2018 McGill Music Conference, Montreal, Canada.
- Reuter, C., Jewanski, J., Saitis, C., Czédik-Eysenberg, I., Siddiq, S., & Oehler, M. (2018b). Colors and timbres—Consistent color-timbre mappings at non-synesthetic individuals. In *Proceedings of the 34th Jahrestagung der Deutschen Gesellschaft für Musikpsychologie: Musik im audiovisuellen Kontext*.
- Ries, H. A. (1969). The elicitation of mediators and colors as a function of stimulus tone frequency. *Dissertation Abstracts International*, 30(6A), 2395.
- Riggs, L. A., & Karwoski, T. (1934). Synaesthesia. *British Journal of Psychology*, 25, 29–41.
- Rimington, A. W. (1895). *A new art: Colour-music*. A paper read at St. James's Hall on June 6, 1895, published in pamphlet form by Messrs. Spottiswoode & Co., New St. Square. June 13, 1895. [Reprinted in 'Colour Music, the Art of Light', by A. B. Klein, Lockwood, London, 1930, pp. 256–261].
- Rodriguez-Pereyra, G. (2002). *Resemblance nominalism: A solution to the problem of universals*. Oxford University Press.
- Ronga, I., Bazzanella, C., Rossi, F., & Iannetti, G. (2012). Linguistic synaesthesia, perceptual synaesthesia, and the interaction between multiple sensory modalities. *Pragmatics and Cognition*, 20(1), 135–167. <https://doi.org/10.1075/pc.20.1.06ron>
- Root, R. T., & Ross, S. (1965). Further validation of subjective scales for loudness and brightness by means of cross-modality matching. *American Journal of Psychology*, 78, 285–289. <https://doi.org/10.2307/1420502>
- Sabaneev, L., & Pring, S. W. (1929). The relation between sound and colour. *Music and Letters*, 10, 266–277. <https://doi.org/10.1093/ml/10.3.266>

- Sabaneyev, L. (1911). *On sound-color accordance*. *Muzyka*, 9 (January 29), 196–200. in Russian.
- Schacter, D. L., Gilbert, D. T., & Wegner, D. M. (Eds.). (2010). *Psychology*, 2nd Edition. New York: Worth Publishers.
- Schauss, A. G. (1981). Application of behavioral photobiology to human aggression: Baker-Miller pink. *International Journal of Biosocial Research*, 2, 25–27.
- Schauss, A. G. (1985). The physiological effect of color on the suppression of human aggression: Research on Baker-Miller Pink. *International Journal of Biosocial Research*, 7(2), 55–64.
- Schier, D. S. (1941). *Louis Bertrand Castel, anti-Newtonian scientist*. Torch Press.
- Schifferstein, H. N. J., & Tanudjaja, I. (2004). Visualizing fragrances through colors: The mediating role of emotions. *Perception*, 33(10), 1249–1266. <https://doi.org/10.1068/p5132>
- Schloss, K. B., Lessard, L., Walmsley, C. S., & Foley, K. (2018). Color inference in visual communication: The meaning of colors in recycling. *Cognitive Research: Principles and Implications*, 3(5). <https://doi.org/10.1186/s41235-018-0090-y>
- Schöffer, N. (1985). Sonic and visual structures. *Leonardo*, 18(2), 59–68. <https://doi.org/10.2307/1577872>
- Schueller, H. M. (1953). Correspondences between music and the sister arts, according to 18th century aesthetic theory. *Journal of Aesthetics and Art Criticism*, 11, 334–359.
- Sebba, R. (1991). Structural correspondence between music and color. *Color Research & Application*, 16, 81–88. <https://doi.org/10.1002/col.5080160206>
- Shen, Y. (1997). Metaphors and conceptual structure. *Poetics*, 25(1), 1–16.
- Shen, Y., & Aisenman, R. (2008). ‘Heard melodies are sweet, but those unheard are sweeter’: Synaesthetic metaphors and cognition. *Language and Literature*, 17(2), 107–121. <https://doi.org/10.1177/0963947007088222>
- Shen, Y., & Cohen, M. (1988). How come silence is sweet but sweetness is not silent: A cognitive account of directionality in poetic synaesthesia. *Language and Literature*, 7(2), 123–140. <https://doi.org/10.1177/096394709800700202>
- Shen, Y., & Gadir, O. (2009). Target and source assignment in synaesthetic possessive constructions. *Journal of Pragmatics*, 41(2), 357–371.
- Shen, Y., & Gil, D. (2007). Sweet fragrances from Indonesia: A universal principle governing directionality in synaesthetic metaphors. In W. van Peer & J. Auracher (Eds.), *New beginnings in literary studies* (pp. 49–71). Cambridge Scholars Publishing.
- Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, 27(2), 125–140.
- Shepard, R. N. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika*, 39(4), 373–421. <https://doi.org/10.1007/BF02291665>
- Shibuya, Y., Nozawa, H., & Kanamaru, T. (2007). Understanding synesthetic expressions: Vision and olfaction with the physiological = psychological model. In M. Plümacher & P. Holz (Eds.), *Speaking of colors and odors* (pp. 203–226). John Benjamins.
- Simmer, J., & Hubbard, E. M. (Eds.). (2013). *The Oxford handbook of synesthesia*. Oxford University Press.
- Simmer, J., Cuskley, C., & Kirby, S. (2010). What sound does that taste? Cross-modal mapping across gustation and audition. *Perception*, 39(4), 553–569. <https://doi.org/10.1068/p6591>
- Simpson, R. H., Quinn, M., & Ausubel, D. P. (1956). Synaesthesia in children: Association of colors with pure tone frequencies. *Journal of Genetic Psychology*, 89, 95–103. <https://doi.org/10.1080/00221325.1956.10532990>
- Sliz, W. (1942). Heine’s synesthesia. *Publications of the Modern Language Association*, 57, 469–488.
- Smith, L. B. (1987). Perceptual relations and perceptual language: A commentary. *Monographs of the Society for Research in Child Development*, 52(215), 94–100.
- Smith, L. B., & Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive Psychology*, 24(1), 99–142.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73, 971–995. <https://doi.org/10.3758/s13414-010-0073-7>
- Spence, C. (2012). Synaesthetic marketing: Cross sensory selling that exploits unusual neural cues is finally coming of age. *The Wired World in 2013*, November, pp. 104–107.
- Spence, C. (2015). Cross-modal perceptual organization. In J. Wagemans (Ed.), *The Oxford handbook of perceptual organization* (pp. 649–664). Oxford University Press.
- Spence, C. (2018). Sensory substitution: Unfulfilled promises & fundamental limitations. In F. Macpherson (Ed.), *Sensory substitution and augmentation* (pp. 251–266). Oxford University Press.
- Spence, C. (2019a). Multisensory experiential wine marketing. *Food Quality & Preference*, 71, 106–116. <https://doi.org/10.1016/j.foodqual.2018.06.010>
- Spence, C. (2019b). On the relative nature of (pitch-based) cross-modal correspondences. *Multisensory Research*, 32(3), 235–265. <https://doi.org/10.1163/22134808-20191407>
- Spence, C. (2020a). Assessing the role of emotional mediation in explaining crossmodal correspondences involving musical stimuli. *Multisensory Research*, 33, 1–29. <https://doi.org/10.1163/22134808-20191469>
- Spence, C. (2020b). Designing scented colours: On the art & science of olfactory-colour crossmodal correspondences. *The Baltic International Yearbook of Cognition, Logic and Communication*, 14(1), 2. <https://doi.org/10.4148/1944-3676.1125>
- Spence, C. (2020c). Temperature-based crossmodal correspondences: Causes & consequences. *Multisensory Research*, 33, 645–682. <https://doi.org/10.1163/22134808-20191494>
- Spence, C. (2020d). Olfactory-colour crossmodal correspondences in art, science, & design. *Cognitive Research: Principles & Implications (CRPI)*, 5(52). <https://doi.org/10.1186/s41235-020-00246-1>
- Spence, C. (2021). Musical scents: On the surprising absence of scented musical/auditory events entertainments, and experiences. *i-Perception*, 12(4), 1–31. <https://doi.org/10.1177/20416695211038747>
- Spence, C. (2022). Searching for perceptual similarity within, and between, the (chemical) senses. *i-Perception*, 13(5), 1–15. <https://doi.org/10.1177/20416695221124154>
- Spence, C., & Bayne, T. (2015). Is consciousness multisensory? In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its modalities* (pp. 95–132). Oxford University Press.
- Spence, C., & Deroy, O. (2012). On the shapes of tastes and flavours. *Petits Propos Culinaires*, 97, 75–108.
- Spence, C., & Deroy, O. (2013). Crossmodal mental imagery. In S. Lacey & R. Lawson (Eds.), *Multisensory imagery: Theory and applications* (pp. 157–183). Springer.
- Spence, C., & Di Stefano, N. (2022a). Crossmodal harmony: Looking for the meaning of harmony beyond hearing. *i-Perception*, 13(1), 1–40. <https://doi.org/10.1177/20416695211073817>
- Spence, C., & Di Stefano, N. (2022b). Coloured hearing, colour music, colour organs, and the search for perceptually meaningful correspondences between colour and sound. *i-Perception*, 13(3), 1–42. <https://doi.org/10.1177/20416695221092802>
- Spence, C., & Di Stefano, N. What, if anything, can be considered an amodal sensory dimension? *Psychonomic Bulletin & Review*, Submitted.
- Spence, C., & Levitan, C. A. (2021). Explaining crossmodal correspondences between colours and tastes. *i-Perception*, 12(2), 1–28. <https://doi.org/10.1177/20416695211018223>

- Spence, C., Deroy, O., & Bremner, A. (2013). Questioning the utility of the concept of amodality: Towards a revised framework for understanding crossmodal relations. *Multisensory Research*, 26(Supple), 57.
- Spence, C., & Wang, Q. J. (2015a). Wine & music (I): On the cross-modal matching of wine & music. *Flavour*, 4(34). <https://doi.org/10.1186/s13411-015-0045-x>
- Spence, C., & Wang, Q. J. (2015b). Wine & music (II): Can you taste the music? Modulating the experience of wine through music and sound. *Flavour*, 4(33). <https://doi.org/10.1186/s13411-015-0043-z>
- Spence, C., & Wang, Q. J. (2015c). Wine & music (III): So what if music influences taste? *Flavour*, 4(36). <https://doi.org/10.1186/s13411-015-0046-9>
- Spence, C., Wan, X., Woods, A., Velasco, C., Deng, J., Youssef, J., & Deroy, O. (2015). On tasty colours and colourful tastes? Assessing, explaining, and utilizing crossmodal correspondences between colours and basic tastes. *Flavour*, 4(23). <https://doi.org/10.1186/s13411-015-0033-1>
- Stechow, W. (1953). Problems of structure in some relations between the visual arts and music. *Journal of Aesthetics and Art Criticism*, 11, 324–333.
- Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, 64, 153–181. <https://doi.org/10.1037/h0046162>
- Stevens, S. S. (1966). Matching functions between loudness and ten other continua. *Perception & Psychophysics*, 1, 5–8. <https://doi.org/10.3758/BF03207813>
- Stevens, S. S. (1971). Issues in psychophysical measurement. *Psychological Review*, 78, 426–450. <https://doi.org/10.1037/h0031324>
- Stevens, S. S., & Guirao, M. (1963). Subjective scaling of length and area and the matching of length to loudness and brightness. *Journal of Experimental Psychology*, 66, 177–186. <https://doi.org/10.1037/h0044984>
- Strik-Lievers, F., & Huang, C. R. (2016). A lexicon of perception for the identification of synaesthetic metaphors in corpora. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (pp. 2270–2277).
- Suarez de Mendoza, D. F. (1890). *L'audition colorée. Étude sur les fausses sensations secondaires physiologiques et particulièrement sur les pseudo-sensations de couleurs associées aux perceptions objectives des sons* [Coloured hearing. Study on false physiological secondary sensations and particularly on pseudosensations of colours associated with objective perceptions of sounds]. Paris: Doin.
- Sullivan, J. W. N. (1914). An organ on which color compositions are played. The new art of color music and its mechanism. *Scientific American*, 110(8), 163. <https://doi.org/10.1038/scientificamerican02211914-163>
- Sun, X., Li, X., Ji, L., Han, F., Wang, H., Liu, Y., Chen, Y., Lou, Z., & Li, Z. (2018). An extended research of crossmodal correspondence between color and sound in psychology and cognitive ergonomics. *PeerJ*, 6, e4443. <https://doi.org/10.7717/peerj.4443>
- Taussig, M. (1993). Mimesis and alterity: A particular history of the senses. *Routledge*. <https://doi.org/10.4324/9781315227689>
- The taste organ. (1926). *Science & Invention*, 14(3).
- Triarhou, L. C. (2016). Neuromusicology or musiconeurology? “Omni-art” in Alexander Scriabin as a fount of ideas. *Frontiers in Psychology*, 7, 364.
- Trotta, R., Hajas, D., Camargo-Molina, J. E., Cobden, R., Maggioni, E., & Obrist, M. (2020). Communicating cosmology with multisensory metaphorical experiences. *Journal of Science Communication*, 19(2). <https://doi.org/10.22323/2.19020801>
- Tsur, R. (1992). *Toward a theory of cognitive poetics*. Elsevier Science Publishers.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84(4), 327.
- Ullman, S. (1957). Panchronistic tendencies in synaesthesia. In S. Ullman (Ed.), *The principles of semantics* (pp. 266–289). Blackwell.
- Underwood, B. F. (1893). Association of color with sounds. *Science*, 21(541), 329–330. <https://doi.org/10.1126/science.ns-21.541.329>
- von Erhardt-Siebold, E. (1932). Some inventions of the pre-Romantic period and their influence upon literature. *Englische Studien*, 66(3), 347–363.
- von Helmholtz, H. (1878/1971). *Treatise on physiological optics* (2nd ed.). Dover Publications.
- von Hornbostel, E. M. (1931). Über Geruchshelligkeit [On smell brightness]. *Pflügers Archiv für die Gesamte Physiologie des Menschen und der Tiere*, 227, 517–538. <https://doi.org/10.1007/BF01755351>
- Von Hornbostel, E. (1938). The unity of the senses. In Ellis, W. D. (Ed.), *A source book of Gestalt Psychology* (pp. 210–216). Kegan Paul.
- Wagemans, J. (Ed.). (2015). *The Oxford handbook of perceptual organization*. Oxford: Oxford University Press.
- Walker-Andrews, A. (1994). Taxonomy for intermodal relations. In D. J. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp. 39–56). Erlbaum.
- Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quality. *Trends in Cognitive Sciences*, 7(11), 483–488. <https://doi.org/10.1016/j.tics.2003.09.002>
- Wang, Y., & Zhang, Q. (2015). The affective effect of simple geometric shapes. *Advances in Psychology*, 5(8), 471–480. <https://doi.org/10.12677/ap.2015.5806>
- Ward, J., Moore, S., Thompson-Lake, D., Salih, S., & Beck, B. (2008). The aesthetic appeal of auditory-visual synaesthetic perceptions in people without synaesthesia. *Perception*, 37(8), 1285–1296.
- Waterworth, J. A. (1997). Creativity and sensation: The case for synaesthetic media. *Leonardo*, 30(4), 327–330. <https://doi.org/10.2307/1576481>
- Watson, Q. J., & Gunter, K. L. (2017). Trombones elicit bitter more strongly than do clarinets: A partial replication of three studies of Crisinel and Spence. *Multisensory Research*, 30(3/5), 321–335. <https://doi.org/10.1163/22134808-00002573>
- Wegner-Clemens, K., Malcolm, G. L., & Shomstein, S. (2022). How much is a cow like a meow? A novel database of human judgments of audiovisual semantic relatedness. *Attention, Perception, & Psychophysics*, 84, 1317–1327. <https://doi.org/10.3758/s13414-022-02488-1>
- Wehner, W. L. (1966). The relation between six paintings by Paul Klee and selected musical compositions. *Journal of Research in Music Education*, 14, 220–224.
- Wellek, A. (1927). Die Farbe-Ton-Forschung und ihr erster Kongreß [Colour-sound research and its first congress]. *Zeitschrift für Wissenschaft*, 9, 576–584.
- Wells, A. (1980). Music and visual color: A proposed correlation. *Leonardo*, 13(2), 101–107. <https://doi.org/10.2307/1577978>
- Wender, I. (1968). Intensität und Qualität in der Geruchswahrnehmung [Intensity and quality in olfactory perception]. *Psychologische Forschung*, 32, 244–276.
- Werner, H. (1934). L'Unité des sens [The unity of the senses]. *Journal de Psychologie, Normale et Pathologique*, 31, 190–205.
- Whiteford, K. L., Schloss, K. B., Helwig, N. E., & Palmer, S. E. (2018). Color, music, and emotion: Bach to the blues. *i-Perception*, 9(6), 1–27. <https://doi.org/10.1177/2041669518808535>
- Wicker, F. W. (1968). Mapping the intersensory regions of perceptual space. *American Journal of Psychology*, 81, 178–188. <https://doi.org/10.2307/1421262>
- Wilkins, J. (1680). *Mathematical magick*. London.
- Williams, J. M. (1976). Synesthetic adjectives: A possible law of semantic change. *Language*, 52, 461–478.

- Winter, B. (2011). Pseudoreplication in phonetic research. In *ICPhS* (pp. 2137–2140).
- Winter, B. (2016a). Taste and smell words form an affectively loaded part of the English lexicon. *Language, Cognition and Neuroscience*, 31(8), 975–988.
- Winter, B. (2016b). *The sensory structure of the English lexicon*. PhD thesis, University of California Merced.
- Winter, B. (2019). Synaesthetic metaphors are neither synaesthetic nor metaphorical. In L. J. Speed, C. O'Meara, L. San Roque, & A. Majid (Eds.), *Perception metaphors* (pp. 105–126). John Benjamins.
- Winter, B., Perlman, M., & Majid, A. (2018). Vision dominates in perceptual language: English sensory vocabulary is optimized for usage. *Cognition*, 179, 213–220. <https://doi.org/10.1016/j.cognition.2018.05.008>
- Witztum, E., & Lerner, V. (2016). Alexander Nicolaevich Scriabin (1872–1915): Enlightenment or illness? *Journal of Medical Biography*, 24, 331–338. <https://doi.org/10.1177/0967772014537151>
- Woods, A. T., & Spence, C. (2016). Using single colours and colour pairs to communicate basic tastes. *i-Perception*, 7(4). <https://doi.org/10.1177/2041669516658817>
- Woods, A. T., Marmolejo-Ramos, F., Velasco, C., & Spence, C. (2016). Using single colours and colour pairs to communicate basic tastes II; Foreground-background colour combinations. *i-Perception*, 7(5). <https://doi.org/10.1177/2041669516663750>
- Zhou, Y., & Tse, C.-S. (2020). The taste of emotion: Metaphoric association between taste and valence. *Frontiers in Psychology*, 11(986). <https://doi.org/10.3389/fpsyg.2020.00986>
- Zhou, Y., & Tse, C.-S. (2022). Sweet taste brings happiness, but happiness does not taste sweet: The unidirectionality of taste-emotion metaphoric association. *Journal of Cognitive Psychology*, 34(3), 339–361. <https://doi.org/10.1080/20445911.2021.2020797>
- Zigler, M. J. (1930). Tone shapes: A novel type of synaesthesia. *Journal of General Psychology*, 3, 276–287.
- Zilczer, J. (1987). "Color music": Synaesthesia and nineteenth-century sources for abstract art. *Artibus et Historiae*, 8, 101–126. <https://doi.org/10.2307/1483303>
- Zwicker, E., & Fastl, H. (2013). *Psychoacoustics: Facts and models*. Springer Science & Business Media.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.