

# A new Molyneux's problem: Sounds, shapes and arbitrary crossmodal correspondences

Ophelia DEROUY<sup>a1</sup>, and Malika AUVRAY<sup>b</sup>

<sup>a</sup> Centre for the Study of the Senses, Institute of Philosophy,  
University of London, UK

<sup>b</sup> LIMSI, CNRS, Orsay, France

**Abstract.** Several studies in cognitive sciences have highlighted the existence of privileged and universal psychological associations between shape attributes, such as angularity, and auditory dimensions, such as pitch. These results add a new puzzle to the list of arbitrary-looking crossmodal matching tendencies whose origin is hard to explain. The puzzle is all the more general in the case of shape that the shapes-sounds correspondences have a wide set of documented effects on perception and behaviour: Sounds can for instance influence the way a certain shape is perceived (Sweeny et al., 2012). In this talk, we suggest that the study of these crossmodal correspondences can be related to the more classical cases of crossmodal transfer of shape between vision and touch documented as part of Molyneux's question, and reveal the role that movement plays as an amodal invariant in explaining the variety of multimodal associations around shape.

**Keywords.** Crossmodal correspondences, Audition, Touch; Molyneux's problem; Amodal invariants

## Introduction: A contemporary version of Molyneux's problem

How do shapes sound? The question does not seem to make sense metaphysically: Shapes are not endowed with auditory properties. In addition, similarities or differences in shapes do not directly correlate with differences in sounds, given that crucial elements such as density, size, and material properties will make similarly shaped objects sound very differently when they are similarly struck. For instance, a small dense sphere might have the same sound as a bigger and less dense cylinder when both are struck in a similar way; and the rich repertoire of drums should convince us that shape is not all that matters to determine how objects sound.

If the question ‘how do shapes sound?’ needs to be dismissed then, a milder version of the question might be more resistant: Supposing that shapes have sounds, what would their sound be? Surprisingly, several studies in cognitive sciences have highlighted convergent and stable responses to this question, and they have shown the existence of privileged psychological associations between shape attributes and auditory dimensions, such as pitch. When asked which of two shapes, one rounded and the other one angular, should be called ‘Takete’ and which one should be called

---

<sup>1</sup> Correspondence to: Ophelia Deroy, Centre for the Study of the Senses, Institute of Philosophy, University of London, Malet Street, London, UK. ophelia.deroy@sas.ac.uk

'Maluma', most participants answer that the angular shape should be 'Takete' and the rounded one, 'Maluma' (Kohler, 1929, 1947; see also Ramachandran & Hubbard, 2001a and figure 1).

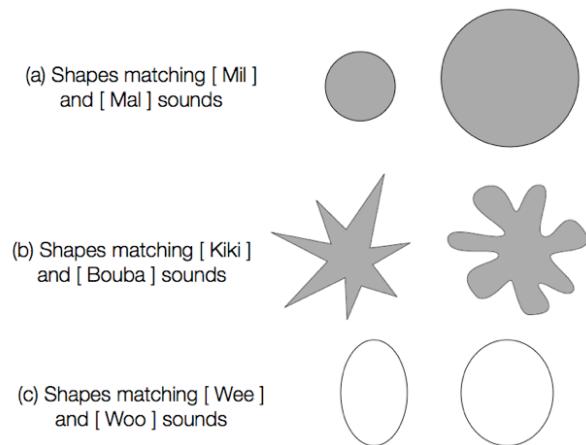


Figure 1. Three examples of crossmodal correspondences, documented between (a) sounds and size by Sapir (1929); (b) sounds and shape (angularity) by Köhler (1929, 1947) and Ramachandran & Hubbard (2001); and (c) sounds and shape (aspect ratio) by Sweeny et al. (2012).

This crossmodal association between shapes and sounds might look surprising at first, but a series of evidence shows it to be present across cultures (Bremner et al., 2013) and from an early age (i.e., four months, see Orztuck et al., 2012, see also Maurer et al., 2006, for evidence in 2 to 2.5 years old). While neurological investigation starts to unveil a specific pattern of neurological activity in the superior / intraparietal regions as well as in frontal areas corresponding to the shape-sound associations (Kovic et al., 2009; Peiffer-Smadja, 2010; see also Bien et al., 2012 for a EEG/ TMS study and Sadaghiani et al., 2009, for a fMRI study of related arbitrary audio-visual correspondences), associations between shapes and sounds is absent in individuals with damage to the angular gyrus (Ramachandran & Hubbard, 2001b), suggesting that this is a robust neuropsychological phenomenon.

What's more, shapes-sounds correspondences have recently been shown to have behavioural consequences, as the visual perception of briefly presented shapes can be affected by certain types of sounds (Sweeny et al., 2012; see also Spence & Deroy, 2012a for a discussion). Sweeny and his colleagues have indeed shown that oval shapes,

whose aspect ratio (relating width to height) varied on a trial-by-trial basis, were rated as looking wider when a /woo/ sound was presented at the same time, and as looking taller when a /wee/ sound was presented instead. By contrast, the perceived shape was not affected by other natural sounds such as birds or engine sounds, showing that a specific crossmodal effect was at stake between these sounds and these shapes. On the one hand, these findings add to a growing body of evidence demonstrating that audio-visual correspondences can have perceptual (as well as decisional) effects (see Parise & Spence, 2012; Deroy & Spence, 2013, for a review). On the other hand, the results concerning sound-shape correspondences add a new puzzle to the list of arbitrary-looking crossmodal matching tendencies whose origin is hard to explain.

The puzzle is all the more general in the case of shape that the shape-sound correspondences have a wide set of documented effects and applications. Besides the aforementioned bias in shape perception, they are shown to facilitate language learning (Imai et al., 2008) and to be exploited in various audio-visual mapping technologies such as music visualization software representing sounds as shapes or sensory substitution devices encoding shapes as sounds (see Deroy & Auvray, 2012).

In the present paper, we suggest that the study of these multimodal associations surrounding shape can be related to the more classical cases of crossmodal transfer of shape between vision and touch documented as part of Molyneux's question (part 2). We review the dominant explanations offered to explain shape-sound correspondences, in terms of conceptual mediation (part 3) and innate hyper-connectivity which is not eliminated by perceptual learning (part 4), before arguing that the hypotheses of associative learning and common neurological representations, proposed to explain the tactile-visual Molyneux's shape transfer can also explain the shape-sound crossmodal matchings. In conclusion, we stress that the hypotheses currently investigated for shape matchings in touch and vision benefit from being extended to the more arbitrary-looking cases of matchings shapes between audition and vision, thereby stressing the multimodal dimension of shape.

## 1. A new Molyneux problem

Arbitrary-looking crossmodal matchings, as they are called (Maurer & Mondloch, 2005; Spence & Deroy, 2012b), can be defined as tendencies to associate distinct sensory features that do not obviously co-occur in experience or in the environment. For instance, moving away from sound-shape pairings for a moment, the tendency to pair higher-pitched sounds with brighter visual surfaces is also shown to be present in adults (Marks, 1974) and in infants (Maurer et al., 2006). So is the tendency to match higher frequency sounds with higher visual locations (e.g. Evans & Treisman, 2010; Spence, 2011 for a review). These pairings occur although brighter objects and animals do not (at least straightforwardly) emit higher pitched sounds than their darker counterparts, and although higher pitched sounds do not regularly come from higher locations in space. The same lack of environmental grounding holds for the correspondence between shapes and sounds: Unless it should turn out that angular objects give rise to sounds that are relevantly different from rounded objects when, for example, they are explored haptically (e.g. Guzman-Martinez et al., 2012), there seems

to be no straightforward environmental correlation between shapes and sounds of objects either.

Crossmodal correspondences between sounds and shapes (or between pitch, brightness and elevation) are difficult to square with the currently popular view that crossmodal associations need to be learned out of the natural multisensory statistics of the environment (see Spence, 2011). Their origin therefore prompts a series of questions. How do such crossmodal correspondences come to be present in humans and other animals? Do they have any ecological value? Determining here whether these sound-shape associations are innate (Ludwig et al., 2011; Maurer & Mondloch, 2005; Maurer et al., 2012) or acquired; and in this case, determining how they are acquired (see Martino & Marks, 1999; Spence, 2011; Walker et al., in press) raise, as we shall see, a new Molyneux's problem, which teaches us new lessons on the multimodal aspect of shapes.

The core of Molyneux's problem, raised initially by Molyneux back in the 17th century, in the heat of the rationalist-empiricist controversies (Locke, 1690; see also Morgan, 1977) is still very much relevant today (e.g., Held et al., 2011, Streri, 2012). The question is to determine whether the crossmodal matching observed between felt and seen shapes at a very early age is acquired through exposure and associative learning, or whether it pre-exists exposure instead. To put it in a philosophical way, the question consists in deciding whether the crossmodal matching of shapes is *a priori* or *a posteriori*. To put it in a psychological way: is the tactile-visual connection for shapes innate / hardwired or acquired?

All past and current replies to Molyneux's problem have been framed on the basis that the matching between tactile and visual shapes targets one and the same environmental property (that is, shape is viewed as an objective or primary quality. Note that Berkeley (1948) is one of the rare philosophers who seems to have accepted that tactile shapes and visual shapes can constitute different objective properties). This objective grounding is what gives the crossmodal matching of tactile shapes and visual shapes a form of necessity and rationality of interest to philosophers.

Now, necessity, rationality and objectivity are what become problematic when we turn to arbitrary crossmodal matchings between sounds and shapes; as they obviously do not target one and the same environmental feature. Certain shapes do not necessarily go with certain sounds. For instance, associating the sound 'Bouba' to a rounded rather than to an angular shape looks irrational and this association does not seem to inform us about an objective regularity. So why would we pair sounds to shapes? Due to these key differences, the mainstream proposals developed for the Molyneux-type of crossmodal associations have not been thought to be relevant to address this question.

The very fact that the crossmodal correspondences between shapes and sounds is called arbitrary comes from the fact that scientists have had a hard time pinning down a regular environmental correlation between the property of being of a certain shape and the property of emitting a certain sound's pitch. Even harder to explain are crossmodal correspondences between shapes and flavours (Deroy & Valentin, 2011) or between symbolic shapes and smells (Seo et al., 2010) which also do not receive a straightforward explanation as internalised statistics of the environments. These other matchings might deserve a separate treatment, but they stress the crux of the problem: If shapes and the other properties are not necessarily or regularly correlated, how could these matchings be learned by association? And if they are not learned by exposure, how could one make sense of the fact that we have evolved to have hard-wired or *a*

*a priori* connections between the representations of shapes and these apparently unrelated properties in our mind / brain?

The competing options that have been recently proposed to explain arbitrary crossmodal matchings between shapes and sounds, as we shall see below, recycle the ones that were once proposed for Molyneux's case but that were subsequently rejected. On the one hand, the idea, initially proposed for Molyneux's cases (see Locke, 1690; or Morgan, 1977 for a review) is that matchings across sensory modalities take place through an association made at the level of ideas or concepts and, on the other hand, the idea, that they are fully present at birth (i.e. that they are *a priori*, see Kant, 1998).

## **2. Sounds-shapes correspondences as conceptually mediated**

The idea that crossmodal matchings require a conceptual mediation is very much the way Molyneux's cases were discussed at the times of Locke and Berkeley, when the connection was supposed to be established between the 'idea' of shape prompted by vision and the 'idea' of shape prompted by touch. The idea of a conceptual mediation is however no longer considered appropriate in order to explain early crossmodal matchings of visual and tactile shapes. However, in the case of arbitrary crossmodal matchings, this hypothesis is pursued by a growing number of researchers: Correspondences between pitch, brightness, and angularity, for instance, have been explained by the cognitive capacity that observers have to represent various sensory features, or dimensions, on a common scale (Martino & Marks, 1999; Walker et al., 2012), to metaphorically map one conceptual domain onto another (Shen, 1997 ; Shen & Eisemann, 2008; Williams, 1976) or to reason analogically (Premack & Premack, 2003; see also Deroy & Spence, 2013; Spence, 2011, for a discussion). Now, the main problem for these conceptual solutions comes from explaining the presence of crossmodal matchings at a very early age (e.g., as early as 4 months, for shapes and sounds, see Orztuck et al., 2012) and the difference between the neurological activations noticed for crossmodal correspondences and semantic or analogical reasoning (Sadhigani et al., 2009).

## **3. Shape-sound correspondences as remnants of non-functional innate connections**

The nativist idea that crossmodal matchings could be present from birth has been eliminated – at least in the case of non-arbitrary matchings – for a long time in favour of the less radically nativist claim that they come from innate learning mechanisms guided by amodal or redundant representations of time, space, and intensity in the brain (see Bahrick & Lickiter, 2012, for a review). The strong nativist option is however still very much present when it comes to explaining arbitrary crossmodal correspondences as shown by the growing popularity of what is called the 'neonatal synesthesia hypothesis' (see Maurer et al., 2012, for a review). The idea here is that these correspondences come from a lack of differentiation of the infant's perceptual apparatus, and persist into adulthood due to a lack of pruning or inhibitory feedback

of some of these non-functional connections (Maurer & Mondloch, 1995; Maurer et al., 2012).

Now, there are good reasons not to go back to strong nativist hypotheses, even to explain the arbitrary crossmodal matchings evidenced in infants. The putative functional role of arbitrary crossmodal matchings as coupling priors in multisensory learning (Ernst, 2007; Spence, 2011) and multisensory integration (Parise & Spence, 2012), or as a kind of crossmodal Gestalt grouping principle (namely, a kind of crossmodal grouping by similarity; see Spence, submitted), together with neurological differences (Sadaghiani et al., 2009; Spence & Parise, in press), are sufficient to distinguish them from non-functional associations that can exist in synaesthetes (no matter whether they are adults or children; see Ward, 2012). This adds to the fact that nativist explanations in general are now hard to support in face of the demand that innate traits are traced back to their genetic encoding (a demand which is not easy to meet for most nativist hypotheses, see Lewkowicz, 2011).

#### **4. Updating the associative learning and common coding hypotheses to explain sound-shape correspondences.**

In this section, we want to argue that the alternative to explain arbitrary crossmodal correspondences either by late conceptual mediation or as being innate is wrongly limited. A first step here consists in stressing that explanations in terms of statistical learning and/or common neural coding have been too swiftly excluded.

The assumption that pairings – between, for instance, angularity and high-pitched sounds – are not regularly experienced by infants is more of an ungrounded assumption. It rather appears to be the default conclusion once one cannot come up with a plausible environmental source for the correlation. It should be more thoroughly investigated by taking into account precise measurements of exposure. Audiovisual correspondences between shapes and sounds might also come from a specific domain, namely speech. The mouth movements observed when someone utters speech sounds like ‘Takete’ or ‘wee’ are more stretched (angular / narrow) than the wider rounded movements observed when one utters ‘Maluma’ or ‘woo’; suggesting a regular correlation between pitch and shapes. This restores the plausibility of an associative learning account, especially compatible with the idea that infants are particularly attentive to face / voice or mouth / sounds in the first months of their life (see the perceptual narrowing hypothesis, Lewkowicz, 2002).

The second assumption that the neurological representation of visual brightness and auditory pitch cannot have anything in common also appears to rely on a predetermined view of what the legitimate common amodal representations in the brain are (i.e., space, time – plus or minus number / magnitude and quantity / intensity; see Marks, 1978). This assumption does not consider other possibilities which are getting investigated in recent work in cognitive neurosciences, that movement (Held et al., 2011) and embodiment could act as common sensibles (note that movement was considered as such by Aristotle and Locke).

Once related to speech, the correspondences between sound and shape can also be explained not merely in terms of audiovisual associations, but also in terms of audio-motor associations, linking the sounds that one hears to the automatic articulatory

movements generated when listening to speech (Galantucci, Fowler, & Turvey, 2006). If the latter account were to be correct, this crossmodal correspondence would then become embodied (Pezzulo et al., 2011), grounded in sensorimotor associations, rather than being based on an external association between two sensory experiences, whose resemblance would be processed in an amodal manner.

One way to distinguish between the statistical and embodied accounts here would be to test whether this correspondence exists only in cases or in species where the vocalising follows the takete-sharp mouth movements rule. Note that this can be contrasted with the correspondence between the sound-size of the source which can be found across species, independently of their rules of vocalization (see Ludwig et al., 2011).

It will further be interesting to determine whether the sound-shape and sound-size crossmodal correspondences are related, and whether the latter has multiple origins (perhaps originating both in external and embodied underlying factors). Understanding the role of embodied vs. external associations would certainly help to link Sweeny et al.'s (2012) results to others showing that the shapes we see - and respond to - can also influence the pitch (or fundamental frequency) of the speech sounds we utter (Parise & Pavani, 2011) or that making a mouth movement (consistent with 'ba' or 'da') can give rise to a McGurk effect (McGurk & MacDonald, 1976) when listening to speech sounds, just as when actually viewing someone else's mouth movements uttering those sounds (see Sams, Mottonen, & Sihvonen, 2005).

## 5. Conclusion

In this concluding section, we want to insist on the importance of focusing on shape-sound correspondences when thinking about shapes, especially in a multi-disciplinary approach. From a global / philosophical perspective, these correspondences encourage a broadening of the investigation of Molyneux's problem, initially focused on tactile and visual shapes, to more contingent associations which can come to matter as much for linguistic and perceptual behaviour. Interestingly, the associative and commonality hypotheses framed here to account for correspondences between shapes and auditory attributes are also at the moment pursued for 'non-arbitrary' matchings of visual and tactile shapes, raising important questions as to how these two shapes might interact, and how situations of single vs. distinct properties can come to differ.

From a more specific and empirical perspective, crossmodal correspondences between shapes and sounds have a role in language acquisition and linguistic intuitions (Imai et al., 2008). They can also explain the use of crossmodal adjectives to talk about sounds (e.g., sharp sounds). But mostly, as we want to highlight, they show all their importance when thinking about the optimization of auditory-visual translations, be it the 'auditory' translation of visual shapes – as in sensory substitution through devices which aim at compensating the loss of sight through a coding / decoding device, such as the vOICe (Meijer, 1992) or the Vibe (Hanneton et al., 2010; see also Auvray & Myin, 2009, for a review) or the visual translation of sounds; for instance as in musical composition software.

## References

- Auvray, M., & Myin, E. (2009). Perception with compensatory devices. From sensory substitution to sensorimotor extension. *Cognitive Science*, 33, 1036-1058.
- Bahrick, L. E., & Lickliter, R. (2012). The role of intersensory redundancy in early perceptual, cognitive, and social development. In A. J. Bremner, D. J. Lewkowicz & C. Spence (Eds.), *Multisensory development* (pp. 183-206). Oxford, UK: Oxford University Press.
- Berkeley, G. (1948–1957). *The Works of George Berkeley, Bishop of Cloyne*. A.A. Luce and T.E. Jessop (eds.). London: Thomas Nelson and Sons. 9 vols.
- Bien, N., ten Oever, S., Goebel, R., & Sack, A. T. (2012). The sound of size: Crossmodal binding in pitch-size synesthesia: A combined TMS, EEG, and psychophysics study. *NeuroImage*, 59, 663-672.
- Bremner, A., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K., & Spence, C. (2013). Bouba and Kiki in Namibia? Western shape-symbolism does not extend to taste in a remote population. *Cognition*, 126, 165–172.
- Deroy, O., & Auvray, M. (2012). Reading the world through sensory substitution devices. *Frontiers in Theoretical and Philosophical Psychology* (DOI: 10.3389/fpsyg.2012.00457).
- Deroy, O., Crisinel, A., & Spence, C. (accepted). Crossmodal correspondences between odours and contingent features: Odours, musical notes, and arbitrary shapes. *Psychonomic Bulletin & Review*.
- Deroy, O., & Spence, C. (2013). Why we are not all synaesthetes (not even weakly so). *Psychonomic Bulletin & Review*, <http://dx.doi.org/10.3758/s13423-013-0387-2>
- Deroy, O., & Valentin, D. (2011). Tasting shapes: Investigating the sensory basis of cross-modal correspondences. *Chemosensory Perception*, 4, 80-90.
- Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision*, 7, 1-14.
- Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, 10, 1-12.
- Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, 13, 361-377.
- Guzman-Martinez, E., Ortega, L., Grabowecky, M., Mossbridge, J., & Suzuki, S. (2012). Interactive coding of visual spatial frequency and auditory amplitude-modulation rate. *Current Biology*, 22, 383-388.
- Hanneton, S., Auvray, M., & Durette, B. (2010). The Vibe: A versatile vision-to-audition sensory substitution device. *Applied Bionics and Biomechanics*, 7, 269-276.
- Held, R., Ostrovsky, Y., de Gelder, B., Gandhi, T., Ganesh, S., Mathur, U., & Sinha, P. (2011). The newly sighted fail to match seen with felt. *Nature Neuroscience*, 14, 551-553.
- Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, 109, 54-65.
- Kant (1998). *Critique of Pure Reason*. Cambridge: Cambridge University Press.
- Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.
- Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. New York: Liveright Publication.
- Kovic, V., Plunkett, K., & Westermann, G. (2009). The shape of words in the brain. *Cognition*, 114, 19-28.

- Lewkowicz, D. J. (2002). Heterogeneity and heterochrony in the development of intersensory perception. *Cognitive Brain Research*, 14, 41-63.
- Lewkowicz, D. J. (2011). The biological implausibility of the nature-nurture dichotomy and what it means for the study of infancy. *Infancy*, 16, 331-367.
- Locke, J. (1690). *An essay concerning human understanding*. London.
- Ludwig, V. U., Adachi, I., & Matzuzawa, T. (2011). Visuo-auditory mappings between high luminance and high pitch are shared by chimpanzees (*Pan troglodytes*) and humans. *Proceedings of the National Academy of Sciences USA*, 108, 20661-20665.
- Meijer, P. B. L. (1992). An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering*, 39, 112-121.
- Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch and loudness. *American Journal of Psychology*, 87, 173-188.
- Marks, L.E. (1978). *The unity of the senses: Interrelations among the modalities*. New York: Academic Press.
- Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception*, 28, 903-923.
- Martino, G., & Marks, L. E. (2001). Synesthesia: Strong and weak. *Current Directions in Psychological Science*, 10, 61-65.
- Maurer, D., & Mondloch, C. J. (2005). Neonatal synesthesia: A reevaluation. In L. C. Robertson & N. Sagiv (Eds.), *Synesthesia: Perspectives from cognitive neuroscience* (pp. 193-213). Oxford: Oxford University Press.
- Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound-shape correspondences in toddlers and adults. *Developmental Science*, 9, 316-322.
- Maurer, D., Gibson, L.C., & Spector, F. (2012). Infant synesthesia. In A. J. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 239-250). Oxford, UK: Oxford University Press.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices, *Nature*, 264 (5588), 746-748.
- Morgan, M. (1977). *Molyneux's question*. Cambridge: Cambridge University Press.
- Ozturk, O., Krehm, M., & Vouloumanos, A. (2012). Sound symbolism in infancy: Evidence for sound-shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*. DOI 2155/10.1016/j.jecp.2012.05.004.
- Parise, C., & Pavani, F. (2011). Evidence of sound symbolism in simple vocalizations. *Experimental Brain Research*, 214, 373-380.
- Parise, C., & Spence, C. (in press). Audiovisual correspondences in the general population. In Simner, J. (ed.) *Oxford Handbook of Synesthesia*. Oxford: Oxford University Press.
- Peiffer-Smadja, N. (2010). Exploring the bouba/kiki effect: A behavioral and fMRI study. Unpublished Ms Thesis, Universite Paris V – Descartes, France.
- Pezzulo, G., Barsalou, L.W., Cangelosi, A., Fischer, M.H., McRae, K., & Spivey, M.J. (2011). The mechanics of embodiment: A dialog on embodiment and computational modelling. *Frontiers in Psychology*, 2, 5.
- Premack, D., & Premack, A. J. (2003). *Original intelligence: Unlocking the mystery of who we are*. New York: McGraw-Hill.
- Ramachandran, V. S., & Hubbard, E. M. (2001a). Synesthesia: A window into perception, thought and language. *Journal of Consciousness Studies*, 8, 3-34.

- Ramachandran, V. S., & Hubbard, E. M. (2001b). Psychophysical investigations into the neural basis of synesthesia. *Proceedings of the Royal Society London B*, 268, 979-983.
- Sadaghiani, S., Maier, J. X., & Noppeney, U. (2009). Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. *Journal of Neuroscience*, 29, 6490-6499.
- Sams, M., Mottonen, R., & Sihvonen, T. (2005). Seeing and hearing others and oneself talk, *Cognitive Brain Research*, 23, 429-435.
- Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, 12, 225-239.
- Seo, H.-S., Arshamian, A., Schemmer, K., Scheer, I., Sander, T., Ritter, G., & Hummel, T. (2010). Cross-modal integration between odors and abstract symbols. *Neuroscience Letters*, 478, 175-178.
- Shen, Y. (1997). Cognitive constraints on poetic figures. *Cognitive Linguistics*, 8, 33-71.
- Shen, Y., & Eisenmann, R. (2008). Heard melodies are sweet, but those unheard are sweeter: Synesthesia and cognition. *Language and Literature*, 17, 101-121.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73, 971-995.
- Spence, C., & Deroy, O. (2012a). Hearing mouth shapes: Sound symbolism and the reverse McGurk effect. *i-Perception*, 3, 550-552.
- Spence, C. & Deroy, O. (2012b) Crossmodal correspondences: Innate or learned? *i-Perception*, 3, 316-318.
- Spence, C., & Parise, C. V. (2012). The cognitive neuroscience of crossmodal correspondences. *i- Perception*, 3, 410-412.
- Streri, A. (2012). Crossmodal interactions in the human newborn: New answers to Molyneux's question. In A. J. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 88-112). Oxford, UK: Oxford University Press.
- Sweeny, T. D., Guzman-Martinez, E., Ortega, L., Grabowecky, M., & Suzuki, S. (2012). Sounds exaggerate visual shape. *Cognition*, 124, 194-200.
- Walker, L., Walker, P., Francis, B. (in press). A common scheme for cross-sensory correspondences. *Perception*.
- Ward, J. (2012). Synesthesia. In B. E. Stein (Ed.), *The new handbook of multisensory processes* (pp. 319-333). Cambridge, MA: MIT Press.
- Williams, J.M. (1976). Synesthetic adjectives: A possible law of semantic change. *Language*, 52, 461-478.