7 Synesthetic Correspondence: An Overview

Lihan Chen

Abstract

Intramodal and cross-modal perceptual grouping based on the spatial proximity and temporal closeness between multiple sensory stimuli, as an operational principle has built a coherent and meaningful representation of the multisensory event/object. To implement and investigate the cross-modal perceptual grouping, researchers have employed excellent paradigms of spatial/temporal ventriloquism and cross-modal dynamic capture and have revealed the conditional constraints as well as the functional facilitations among various correspondence of sensory properties, with featured behavioral evidence, computational framework as well as brain oscillation patterns. Typically, synesthetic correspondence as a special type of cross-modal correspondence can shape the efficiency and effect-size of cross-modal interaction. For example, factors such as pitch/loudness in the auditory dimension with size/brightness in the visual dimension could modulate the strength of the cross-modal temporal capture. The empirical behavioral findings, as well as psychophysical and neurophysiological evidence to address the cross-modal perceptual grouping and synesthetic correspondence, were summarized in this review. Finally, the potential applications (such as artificial synesthesia device) and how synesthetic correspondence interface with semantics (sensory linguistics), as well as the promising research questions in this field have been discussed.

Keywords

Cross-modal · Perceptual grouping · Synesthetic correspondence · Individual differences · Sensory linguistics

L. Chen (✉)
School of Psychological and Cognitive Sciences, Peking University, Beijing, China

Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China

Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China

National Key Laboratory of General Artificial Intelligence, Peking University, Beijing, China

National Engineering Laboratory for Big Data Analysis and Applications, Peking University, Beijing, China
e-mail: CLH@pku.edu.cn

7.1 Introduction

Cross-modal integration entails the correspondence between different sensory properties to reach a coherent and meaningful representation of the environment as well as the target events/objects (Ernst and Bülthoff 2004; Stein and Meredith 1993). Cross-modal correspondences

Y. Gu, A. Zaidel (eds.), *Advances of Multisensory Integration in the Brain*, Advances in Experimental Medicine and Biology 1437, https://doi.org/10.1007/978-981-99-7611-9_7

(CMC) are defined as pairs of associations between two sensory or cognitive processes that are generally agreed upon by most individuals within a population (Brang et al. 2013). In a seminal tutorial by Charles Spence (Spence 2011), CMC has been defined to contain three forms—statistical correspondences, structural correspondence, and semantically mediated correspondences. The three forms have been examined and materialized in a number of studies (Bien et al. 2012; Bremner et al. 2013; Guo et al. 2017; Ngo et al. 2013; Parise and Spence 2012; Piqueras-Fiszman and Spence 2011; Spence and Deroy 2013; Wan et al. 2014; Zeljko et al. 2019; Zhang and Chen 2016). Although the seminal categorization of CMC has been proposed almost a decade ago, the empirical research evidence as well as the concepts developed over the years, could still be nicely fit into this framework. Accumulated evidence has shown the perceptual grouping and intersensory binding of sensory properties indeed affect the ultimate outcome of multisensory integration. The behavioral manifestation and perceptual benefits (as well as interferences with incongruent CMC) are solid from empirical experiments though the underlying mechanism of CMC is far from clear.

Among the kaleidoscope of CMC, the synesthetic correspondence is the most intriguing type and modulates the efficiency and effect-size of cross-modal interaction (Bien et al. 2012; Chen et al. 2016; Grossenbacher and Lovelace 2001; Hidaka and Yaguchi 2018; Parise and Spence 2009; Saenz and Koch 2008). Below I address respectively, the manifestations of synesthetic correspondence, the underlying neurocognitive mechanisms, sound symbolism as an example and the individual differences research approach, as well as the potential applications in this field.

7.2 Synesthetic Correspondence and Cross-modal Integration

The correspondence of properties among different sensory modalities raises two prominent questions as well as challenges. On the one hand, how to compare and measure the qualia (properties) associated with different sensory modalities. For example, as we learn, the properties for different sensory stimuli (e.g., sound pitch vs. visual size) are not comparable apparently, so that how to standardize them to reach "correspondence"? On the other hand, it pits the challenge to describe this correspondence with appropriate words—both with overt/precise words and with covert/implicit descriptions. Linguistic sensory words are usually deficient to describe and capture the large volumes of sensory properties. "The word we use for things like the redness of a rose, beauty, leadership, and charisma. We usually know exactly what we mean, but we have no words to describe let alone define it" (Ross 2018).

The concept of "correspondence" has originally arisen and henceforth extensively exploited in unimodal—visual domain. In the dynamic visual scene, during the continuous eye movements for visual perception, an outstanding question is that how the visual system could establish and maintain the identity of an object under above synergistic "interaction" between environment and observers (say, juggling), i.e., *correspondence problem* (Ullman 1979). Studies using motion correspondence have explored the modulating factors including featural aspect. In Hein and Cavanagh (2012) study, they used Ternus display (a form of ambiguous visual apparent motion) and found contrast polarity, spatial frequency, and size modulate the perceived motion categorization (such as bias in element motion), and hence demonstrated a spatiotopic-based feature bias (Hein and Cavanagh 2012). The Ternus display has been exploited to investigate the perceptual grouping in auditory modality (Wang et al. 2014), tactile modality (Chen et al. 2010), and multisensory processing (Fig. 7.1) (Chen et al. 2018).

Cross-modal correspondences (CMCs) have been defined as "a tendency for a feature, attribute, dimension, or stimulus in one sensory modality, either physically present or merely imagined, to be matched (or associated) with a feature, attribute, dimension, or stimulus in another sensory modality" (Parise 2016). Usually, audiovisual correspondences have been adopted to investigate CMC. Below, synesthetic corre-

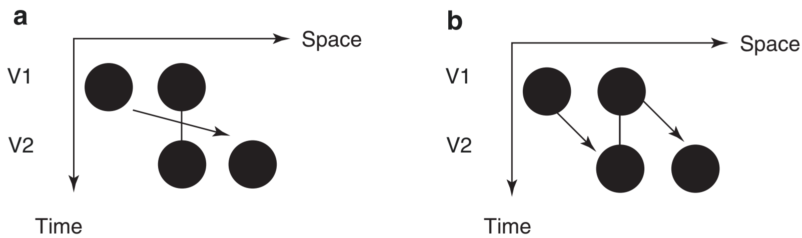

Fig. 7.1 Ternus display and stimulus configurations. Visual Ternus display contained two types of exclusive percepts: (**a, b**). "Group" motion in which the two frames move from one side to the other as a whole. (From Fig. 1, Chen et al. (2018), Journal of Experimental Psychology: General, with permission)

spondence was illustrated with appropriate examples following the categorization proposed by Spence (2011).

The "structural" correspondence, has been classically materialized in spatial/temporal ventriloquism paradigm, in which the perceived temporal onset has been shifted due to the presence of concurrent task-irrelevant distractors (temporal ventriloquism) or been mislocated due to the proximate distractors (spatial ventriloquism) (Bertelson and Aschersleben 2003; Bertelson et al. 2000; Chen and Vroomen 2013; Morein-Zamir et al. 2003; Ogulmus et al. 2018; Orchard-Mills et al. 2016; Slutsky and Recanzone 2001; Tian et al. 2020; Vroomen et al. 2004; Vroomen and Keetels 2006). According to this structural account, with the ventriloquism paradigm, perceptual grouping across space and time is operating in an automatic fashion to arrange stimuli according to their magnitudes/intensities. Or put in another way, the correspondence modulates the spatiotemporal factors in intra- and cross-modal groupings. For instance, Parise and Spence (2009) once reported that spatial and temporal ventriloquism effects are enhanced for cross-modally congruent pairs of auditory and visual stimuli as compared with pairs of incongruent stimuli (Parise and Spence 2009), prominently, pitch-size cross-modal correspondence modulated the spatial ventriloquism effect as well (Parise and Spence 2008). As a counterpart for spatial ventriloquism, Orchard-Mills et al. (2016) investigated whether cross-modal correspondence between auditory pitch and visual elevation modulates temporal ventriloquism. Using temporal order judgment, they asked participants to judge the order of two visual stimuli (above and below fixation) across a range of stimulus onset asynchronies (SOAs). The results show that incongruent pairings between pitch and elevation abolish temporal ventriloquism. In contrast, the potential "congruent" ("facilitation") effect was dependent on the saliency of the cross-modal mapping (Orchard-Mills et al. 2016).

In statistical account, cross-modal correspondences can be thought of as the internalization of the statistical regularities of the stimuli associations (combinations) and the environment. Human observers use the similarity in the temporal structure of multisensory signals to solve the correspondence problem—inferring causation from correlation (Parise et al. 2012). With an elegant design, Parise et al. (2012) uncovered the role of correlation between the fine temporal structure of auditory and visual signals in causal inference (i.e., correspondence problem). They found that in a localization task with visual, auditory, and combined audiovisual targets, the improvement in precision for combined (auditory and visual stimuli) relative to unimodal targets (only with auditory stimuli) was statistically optimal only when audiovisual signals were correlated. In a following study, Parise and Ernst further generalized the research scope and proposed that "correlation" detection as a general mechanism for multisensory integration. They termed the multisensory correlation detector which integrates related multisensory signals through a set of temporal filters followed by linear combination. The "correlation" detection mechanism successfully explains a range of phenomena with physiologically plausible model

(Parise and Ernst 2016). The above two representative studies suggested that the "statistical" relations between signals from two different modalities, have essentially built "correspondence" in temporal relations to facilitate perceptual decision-making, such as spatial localization for auditory target.

In cross-modal interaction scenario, Zhang and Chen (2016) applied the "correspondence" principle and manipulated the statistical relationship between the temporal onsets of two beeps and two visual Ternus frames, in which the visual frames with different colors ("red" or "black") were associated with certain audiovisual temporal structures (Fig. 7.2). They found that by binding the relevant temporal information and stimuli properties—through manipulating the probabilities of the occurrences of audiovisual events, the perception of visual Ternus motion could be quickly recalibrated. Observers could implicitly employ the temporal (interval) relations between the target events as a prior and demonstrate the

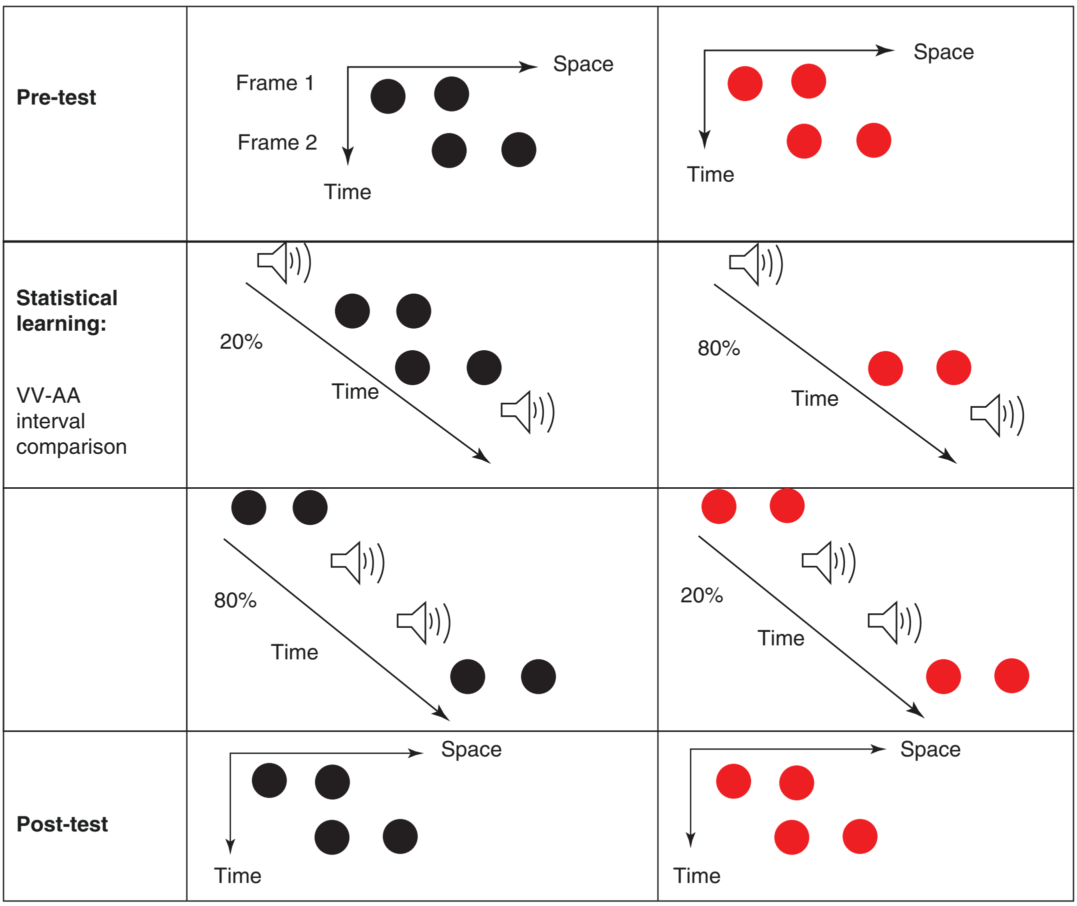

Fig. 7.2 Ternus displays (for pretest and posttest) and illustrations of the stimuli configurations for Experiment 1 in Zhang and Chen (2016). Two kinds of Ternus display were adopted (with black frames or red frames). In the training session, for black-black (BB) configuration, in 80% trials, first sound preceded the first red visual frame and the second sound trailed the second red visual frame by 80 ms (i.e., "AVVA" condition). In red-red (RR) configuration, in 80% trials, the first visual frame preceded the first beep and the second visual frame trailed the second beep by 80 ms (hereafter "VAAV" condition). The inter-stimulus-interval (ISI) between two Ternus frames were randomly set from 50 to 230 ms. For the rest 20% trials, the RR and BB configurations were associated with temporal structures of AVVA and VAAV, respectively. (From Fig. 2 in Zhang and Chen (2016), Frontiers in Psychology, with permission)

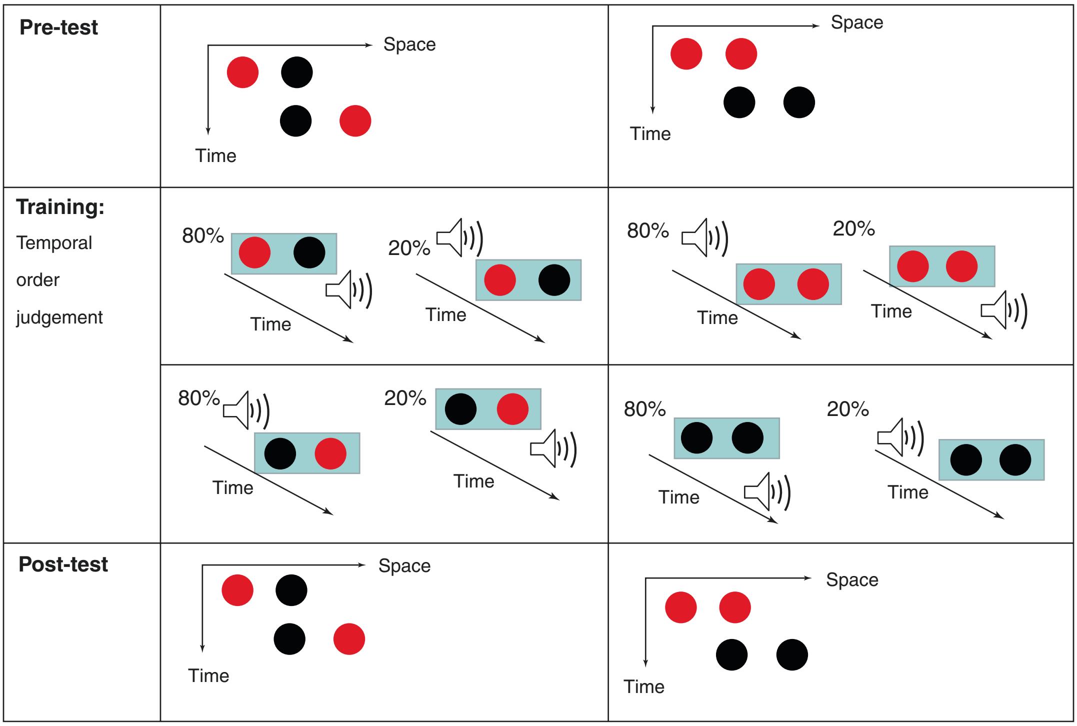

Fig. 7.3 Ternus displays (for pretest and posttest) and illustrations of the stimuli configurations for Experiment 2 in Zhang and Chen (2016). Two kinds of Ternus display were adopted. In the training session, for RB-BR configuration, in 80% trials, the RB frames preceded the sound beep and BR frames trailed the beep. In contrast, 20% of RB frames trailed the sound beep and 20% BR frames preceded the sound beep. For RR-BB configuration, 80% of the RR frames trailed the sound beep and BB frames preceded the beep. In contrast, 20% of RR frames preceded the sound beep and BB frames trailed the sound beep (From Fig. 5 in Zhang and Chen (2016), Frontiers in Psychology, with permission)

observable temporal aftereffects which give rise to the featured proportional changes of reporting element motion vs. group motion of Ternus display (Zhang and Chen 2016) (Fig. 7.3).

Synesthetic correspondence represents a natural tendency of associating sensory properties between two typical modalities and contains semantic meanings (values). For example, elevation serves as a fundamental organizational dimension for many cross-modal correspondences. Visual elevation corresponds to the high pitch and while the low visual location (height) naturally corresponds to the low auditory pitch, according to our gained world knowledge. People often show a systematic tendency to associate moving objects with changing pitches.

Saenz and Koch reported that auditory synesthesia does indeed exist with evidence from four healthy adults for whom seeing visual flashes or visual motion automatically causes the perception of sound (Saenz and Koch 2008).

By employing the synesthetic audiovisual correspondences between the visual Ternus movement directions (upward and downward) and the changes of pitches of concurrent glides, Guo et al. (2017) examined how human observers' cognitive abilities, typically, the cognitive styles (field-dependent vs. field-independent) modulate/make distinction the relations between the properties of pitch glides and the perception of visual apparent motion (Fig. 7.4). The results indicated that for pitch ascending (decreasing) glides, when they were paired with moving upward (downward) visual Ternus, observers would perceive dominant percept of "element motion." Importantly, field-independent observ-

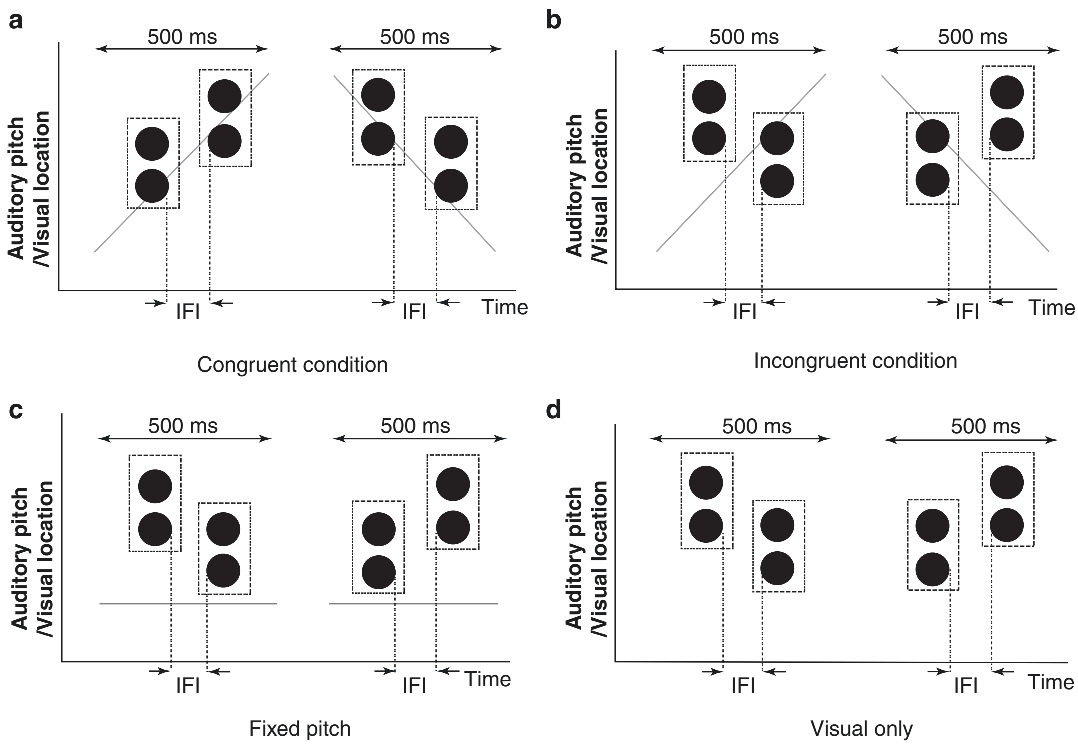

Fig. 7.4 An illustration of the stimuli configuration for Guo et al. (2017). There were four configurations: "Congruent," "Incongruent," "Fixed-pitched," and "Visual only." The former three conditions contained audiovisual stimuli. With the differential directions of the solid lines, the glide was either one of the three types: ascending, descending, or invariant. "Visual only" referred to visual Ternus display without accompanying sounds. The durations of audiovisual stimuli were constant at 500 ms. For "congruent" configuration, Ternus displays with upward/downward motion were given together with ascending-pitched/descending-pitched glides. For "incongruent" condition, upward (downward) Ternus motion was synchronized with descending (ascending) pitch glide. For "fixed-pitched" condition, a glide with pure tone (500 Hz) was given. IFI = inter-frame-interval. (From Fig. 2 in Guo et al. (2017), Multisensory Research, with permission)

ers tended to readily identify the audiovisual events and use the principle of cross-modal correspondence to demonstrate the above "biased" perception (Guo et al. 2017).

Synesthetic correspondence with semantic (linguistic) relations is a higher level and abstract process. Often stimuli may be related because of the similar terms we sometimes use to describe stimuli presented in different senses, such as the "high"/"low" could be used as common words for both pitch and spatial elevation in English (also in Chinese) (Martino and Marks 1999; Sadaghiani et al. 2009). One typical example of semantic correspondence is sound symbolism (addressed in detail in the following session), referring to the non-acoustic information (literal meaning) of auditory property. The "higher-level" nature is supported by the evidence that the sensitivities to sound symbolism increased with acquired (greater) language experience.

7.3 Sound Symbolism as a Form of Synesthetic Correspondence

According to the Cassierer's framework, language provides two very different ways of expressing interrelationships among the senses. One is analogical and the other is symbolic. Language serves as a formal representation to communicate cross-modal equivalence of (sensory) properties (Cassirer 1953). In sensory linguistic, sound symbolism is not referring to onomatopoeic expressions like words—"buzz," but reflects some nonacoustical property of nature. The relationship between sound and meaning is indeed intrinsic. The most significant and popular exemplar of sound symbolism is the "bouba-kiki" effect, which refers to the *shape-sound* correspondence (Barton and Halberstadt 2018; Cuskley et al. 2017; De Carolis et al. 2018; Fort et al. 2015; Fryer et al. 2014; Gold and Segal 2017; Karthikeyan et al. 2016; Köhler 1929; Maurer et al. 2006; Peiffer-Smadja and Cohen 2019; Sakamoto and Watanabe 2018). People usually associate "Bouba" to an image of a rounded object and "Kiki" to an image of a spiky object. Indeed, features, phonemes, and letters may all contribute to sound symbolism. Sound symbolism can be found in different literature forms and across different levels.

In his seminal book "The Unity of the Senses", Lawrence E. Marks introduced a very interesting topic for Sound Symbolism in Poetry, in which he addressed the expressions of sensory analogy and correspondence in literature, and mainly with examples from the poetry of the nineteenth century in the western world (Marks 2014). Poetry is a container that showcases the sound symbolism in speech. Sounds suggest meanings through suprasensory perceptual attributes and one way that sounds symbolize meaning is by means of affect. As reported by earlier studies, /a/ and /o/ were judged more pleasant than /u/ (Roblee and Washburn 1912). Speech sounds could represent more general and delicate sensory properties such as the gradations in brightness and changes in spatial size. Macdermott (1940) summarized how low-pitched and high-pitched vowels correspond to the sensory attributes by the method of semantic differentiation. On the other hand, synesthetic metaphor conveys intrinsic cross-modal relationship and suggests a trend of the evolution for semantic (language) adjectives that applied to typical sensations, such as "sharp" and "dull" for descriptions of tactile and visual sensations (Macdermott 1940). Westbury et al. (2018) constructed graphic representation of correlations between hybrid model estimates (adjectives) of the categories, revealing the overlapping of estimations to different degrees. Therefore, sound symbolism operates at broad categorical levels (Westbury et al. 2018).

7.4 Neurocognitive Mechanism of Synesthetic Correspondence

Researchers have developed experimental paradigms of explicit (cross-modal matching, including forced-choice) (Kohler 1947) and implicit associations (with reaction times) (Hung et al. 2017; Parise and Spence 2012; Westbury 2005) as well as Transcranial Magnetic Stimulation (TMS) (Bien et al. 2012), to address the neural substrates which are responsible for the (synesthetic) correspondence.

Bien et al. (2012) had carried out a representative study to uncover the cross-modal binding in pitch-size synesthesia by adopting a combined TMS, EEG, and psychophysics approach. Using an auditory spatial localization task (Ventriloquist effect), they found that the performance of localization was impaired with the congruent mapping of pitch-size. Synesthetic congruency results in multisensory binding, and consequently auditory sources are misallocated to the visual source. To pin down the neural substrates, they further purposely disturbed the right intraparietal sulcus (via TMS) and observed the ventriloquist effect was abolished and largely identified right intraparietal sulcus being the critical cerebral area to account for the pitch-size mapping. Further correlation analysis (together with ERP) indicated the origin of synesthetic pitch-size mappings to a right intraparietal involvement around 250 ms (Bien et al. 2012).

Synesthetic correspondence is linked with the adjacency of the neural anatomies between two brain areas, which are responsible for encoding/representing the individual (two) sensory properties, respectively. An interesting question is whether there is functional dominant of a sensory modality over the other as in typical multisensory interaction. Evidence has shown that pitch-height associations are vision-dependent, suggesting the visual experience plays an important role in the synesthetic correspondence (Deroy et al. 2016a, b; Fryer et al. 2014). The dependency on dominant sensory modality (visual modality) in synesthetic correspondence has been studied with special group who have deficiency in one sensory modality (such as vision). For example, Bottini et al. (2019) revealed an intriguing possibility that people with early blindness may be more sensitive to (some) iconic features of language compared to sighted (Bottini et al. 2019). In the absence of a direct visual stimulus, visual imagery plays a role in cross-modal integration. As an important replication of "Bouba-Kiki" effect that has been demonstrated in visual-auditory modalities over 80 years ago, Fryer et al. (2014) implemented "Bouba-Kiki" effect in auditory-haptic modalities in which participants received "felt" shapes but did not see them. The sighted participants showed a robust "Bouba-Kiki" effect while people with visual impairments had less pronounced effect.

The neural mechanisms for synesthesia provide valuable information for the corresponding neural mechanism of synesthetic correspondence since synesthesia has been recognized as a specialized/extreme form of "synesthetic correspondence" in its wide spectrum (Marks and Mulvenna 2013). The above cross-talk between two brain areas suggests one representative hypothesis regarding synesthesia: direct cross-activation. The cross-activation model suggests that synesthesia has been attributed to specific anatomical pathways, which are weak or absent for non-synesthetes. Those pathways provide neural underpinnings that a direct link between one sensory modality to another becomes viable (Hubbard et al. 2005, 2011; Ramachandran and Hubbard 2001a, b, 2003).

The hyperbinding account, however, assumes common brain areas for synesthetes and non-synesthetes. However, the two groups differed in the function as well as efficiency of the cross-talk between two senses. Especially, the synesthetes generally have weaker inhibition of the feedback from the inducers (Grossenbacher and Lovelace 2001).

7.5 Cognitive Benefits of Synesthetic Associations and the Role of Training (Experience)

Synesthetic associations bring forth cognitive benefits or differentiations in cognitive performance, including executive control functions, number-space perceptions, and memory functions. Synesthesia broadly impacts perception but has differential impacts (McCarthy and Caplovitz 2014). Rouw et al. (2013) investigated the relationship between synesthesia and executive control functions. Using traditional executive control paradigms and Stroop tasks, they suggested no clear relationship between executive control functions and synesthetic behavioral effects (Rouw et al. 2013).

Gertner et al. (2013) found that synesthetes' automaticity of processing given semantic meaning of numerals were affected by their number-space perceptions. Specifically, both synesthetes and their control peer exhibit the similar, classic size congruency effect (SiCE) for numerical block. However, for the case of physical block, synesthetes demonstrated weak automaticity in processing numerical magnitude when the numbers for comparison were placed incompatibly concerning their relative position in the format of "synesthetic" number space (Gertner et al. 2013).

Ramachandra (2016) investigated the influence of lexical-gustatory synesthesia on memory, using a paired-associate learning task. A single subject (synesthete) predicted that her learning would be better in the "taste" condition when compared to the "no taste" condition, showing a "foresight bias" in the vein of metamemory task (judgement-of-learning) although no significant differences emerged between the "no taste" and "taste" conditions (Ramachandra 2016).

Synesthetic associations are experience dependent. Using sound symbolic novel words and foreign words meaning round and pointy, and conducting a cross-modal matching task (i.e., word-referent mapping, picking round/pointy picture to correspond given word), Tzeng et al. found that while 3-year-old participants showed chance performance, 5- and 7-year-olds showed reliably above-chance performance. In this means, sound symbolism sensitivity is a manifestation of the pre-linguistically available cross-modal mappings (Tzeng et al. 2017). Intersensory correspondences are based on neural connections present very early in life. However, the correspondences are dependent on the specific combinations of cross-modal properties pair.

Haryu and Kajikawa (2012) found 10-month-olds paired a higher frequency tone with an object of a brighter color and associated a lower frequency tone with darker color. However, for the size dimension, the same group did not always pair a higher frequency tone with a smaller object, nor did they pair a lower frequency tone with a larger object. Those findings indicated infants have an initial tendency to match pitch with brightness, while the pitch-size correspondences were loosely connected but could be formed tightly through statistical learning of the simultaneously presented stimuli pair ("pitch-size") (Haryu and Kajikawa 2012).

Extensive cognitive training has been shown to generate synesthesia-like phenomenology. To address the question that whether these experiences are accompanied by neurophysiological changes characteristic of synesthesia will provide a unique opportunity to elucidate the neural basis of perceptual plasticity relevant to conscious experiences. Rothen et al. (2018) confirmed that overtraining synesthetic associations results in synesthetic phenomenology, by using Stroop tasks to reveal synesthesia-like performance following training (Rothen et al. 2018). Goodhew et al. (2015) found that synesthetes produce substantially greater semantic priming magnitudes (using a lexical decision task and a semantic categorization task), unrelated to their specific synesthetic experience (Goodhew et al. 2015). van Petersen et al. (2020) found individuals with sequence-space synesthesia (SSS) who perceive sequences like months, days, and numbers in certain spatial arrangements, have enhanced spatial navigation skills, in addition to the established cognitive benefits such as enhanced mental rotation, more vivid visual imagery and an advantage in spatial processing (van Petersen et al. 2020).

Synesthetic correspondence entails the cognitive ability that appropriately binds the elementary properties of visual and auditory events, the weak correspondence might exist in atypical developing groups (including dyslexics and autistic individuals).

For a starter in reading, they should initially form a basic correspondence between orthographic tokens and phonemic utterances. The associations between letters and sounds remained intact for dyslexics (Blomert 2011). —Synesthetic correspondence, as a type of cross-modal association, has been revealed to remain robustly for dyslexic children—match larger visual shapes with lower auditory pitch, or smaller visual forms with higher auditory pitches (Chen et al. 2016). With the visual temporal order judgment (TOJ) task (two consecutively presented discs were enclosed by two beeps), Chen et al. (2016) also found that the congruent audiovisual pair boosted the TOJ performance by reducing the just noticeable difference (JND), whereas the incongruent audiovisual pair degraded the performance by increasing the JND. The modulation magnitudes were larger in dyslexic children than in normal peer. Importantly, this very basic perceptual performance of TOJ, has been revealed to be correlated with the higher order cognitive expertise such as reading skills (measured by scores of character recognition test and reading fluency) (Chen et al. 2016).

Occelli et al. (2012) reported that when children with Autism Spectrum Disorders (ASDs, 6–15 years) were asked to indicate which of two bouncing visual patterns was making a centrally located sound. In this parametric design, the visual size, surface brightness as well as visual shape were manipulated; the sounds were modified in pitches. The performance of ASDs were at chance levels, indicating that the general poor capabilities to integrate auditory and visual inputs, might constrain the synesthetic associations.

7.6 Synesthetic Correspondence in Sensory Branding

Assigning the right name to a given product is extremely important, as it can help enhance the overall product experience, the product satisfaction as well as ensure the subsequent, repeated consumption (Spence 2012). Sound symbolism has been exploited extensively in sensory branding field for the last decade. In a range of commercial fruit pulps/juices, Ngo et al. (2013) examined shape and sound symbolism. British and Colombian participants associate sweet juices with lower sourness to rounder shapes, while they consistently map sour ones to angular shapes (Ngo et al. 2013). Those findings suggested that in packaging and labeling, it is important to obey these correspondences, especially when the customers are unfamiliar with those products. A "crossmodal-cognitive" perception model of uniqueness has been developed to account for the appropriate food names in terms of sound symbolism (Favalli et al. 2013). Within the evaluation procedure, the investigation of the product and its definition was composed of the combination of a cross-modal and cognitive approach. The moderating influence of the sound symbolism has been recently discovered in willingness-to-pay (WTP) behavior, in which back vowels and voiced consonants (linked with a larger size of retail store) would lead to higher WTP (Ketron and Spears 2019).

As suggested in the former section, the synesthetic correspondence is experience-dependent. However, the evidence for the cultural variation in synesthetic correspondence is rare. Interestingly, the "Bouba-Kiki" effect has also been observed in the Himba of Northern Namibia, a remote population with little exposure to Western cultural influences, though the effect has been robustly demonstrated in Western participants. However, it is not consistent with the mapping of shape to taste (Himba-associated carbonation to angular shape rather than rounded shape). This finding indicates both general and special (individualized) pattern for cross-modal correspondences under different cultural contexts (Bremner et al. 2013).

Investigations have been made as regards with the role of personality in synesthetic correspondences (or synesthesia). Generally, the synesthetes showed increased intelligence as compared with matched non-synesthetes. This was a general effect rather than bounded to a specific cognitive domain or to a specific (synesthesia-type to stimulus material) relationship. By tapping on the "big five" personality (the NEO-PI-R personality inventory), the expected effect of increased "Openness" in synesthesia was obtained, as well as two unexpected effects in personality traits (increased "Neuroticism" and decreased "Conscientiousness"). The personality and cognitive characteristics were found related to having synesthesia (in general) rather than to particular synesthesia subtypes. Therefore, this piece of finding supports the existence of a general synesthetic "trait," over the notion of relatively independent "types" of synesthesia (Rouw and Scholte 2016).

7.7 Sensory Linguistics Perspective: Plurality or Ambiguous Meanings

As seen from its literally meaning, "synesthetic correspondence" implies a common basis for sensory comparison. This idea was supported by ATOM (A Theory of Magnitude) model (Bueti and Walsh 2009; Walsh 2003). ATOM assumes a unified, cross-modal magnitude system for all kinds of sensory magnitudes (including numerical quantities) (Winter et al. 2015). Therefore, "magnitude" could be a common currency as well as cognitive quality among different sensory modalities.

However, there is plurality or ambiguous meanings in both lexicon and forms of correspondences. For example, the word of "hot" have at least two differential meanings. Its basic meaning is referring to higher "temperature." The other one could be "spicy" for foods.

Interestingly, Rakova (2003) investigated the neural basis for the two perceptual meanings and found they a had been sub-served by the same underlying neural system (Rakova 2003). The plurality of words has found its root in "shared" sensations among different sensory channels. For instance, "Words denoting taste cannot always be separated clearly from those for feel and smell" (Staniewski 2017). This common sharing could not be based on perceptual dimension but also on emotional one. For example, the associations between music and color are based more on emotional correspondences (e.g., between major mode and happiness) rather than on perceptual correspondences (Palmer et al. 2013).

In a sense of "embodied cognition," different senses diversify in their relatedness to one's own body. The perceptual experiences through vision and hearing have been mainly projected onto objects (see also Shen 2008, p. 302).

Touch, taste, and smell, on the other hand, are argued to be relatively more subjective and more related to one's bodily consciousness. Zhao et al. (2019) examined the mapping directionality tendencies of linguistic synesthesia in Mandarin. They adopted a corpus-based approach and categorized three types of directional tendencies for Mandarin synesthesia: unidirectional biased-directional and bidirectional (Zhao et al. 2019). The directional flow is determined by the sensory dominance of one modality over the other. For instance, mapping from TASTE to HEARING rather than the opposite is typical for "unidirectional" tendency.

This directional flow of sensory mapping indeed constrains the synesthetic correspondence among different cohorts of sensory properties.

On the other hand, efficient codability is a psycholinguistic measure of the relative ease of expressing certain percepts. It is generally thought that sight is the most codable sensory modality. Levinson and Majid (2014) say that "in English at least, it seems generally easier to linguistically code colors than (nonmusical) sounds, sounds than tastes, tastes than smells" (p. 415) (Levinson and Majid 2014). A persistent challenge in mapping odors to names has been named as a "muted sense" and hence it is generally believed that smell is the most ineffable sensory modality. Sight is overall more codable in English, which actually mean that sight has more distinct perceptual qualities that are codable. With that said, it seems an "unbalanced" correspondence between literal terms and the sensory properties. Humans can distinguish between millions of different colors (up to ten million distinct colors (Judd and Wyszecki 1975), but languages generally have only very basic color terms (Andrea 2007; Brent and Paul 1969; Siegfried 2007). Likewise, humans can similarly distinguish thousands of different smells (Agapakis and Tolaas 2012; Yeshurun and Sobel 2010), but there are very few smell words at least in English. This mismatch of words to sensory properties indicates the potential ineffability of fine perceptual details and linguistically impoverished in sensory descriptions (naming) (Agapakis and Tolaas 2012; Yeshurun and Sobel 2010).

7.8 Applications: Synesthesia Device

Human senses could be simulated, digitalized, and even visualized. The emergence of artificial cognitive system and digital technology has made the interactive visualization of human senses viable and benefit the applications of synesthesia device as well. Kim et al. (2019) presented flexible artificial synesthesia electronics that visualize continuous and complicated sounds (Kim et al. 2019). It is by far an excellent example to illustrate the synesthetic correspondence by designing an elegant device. They made an electronic device (FASSEL) which contains a thin composite film of a piezoelectric polymer for sound generation and inorganic electroluminescence (EL) microparticles for direct visualization of input sound signals (Fig. 7.5a). Field-induced EL responded to the source sound wave. The main

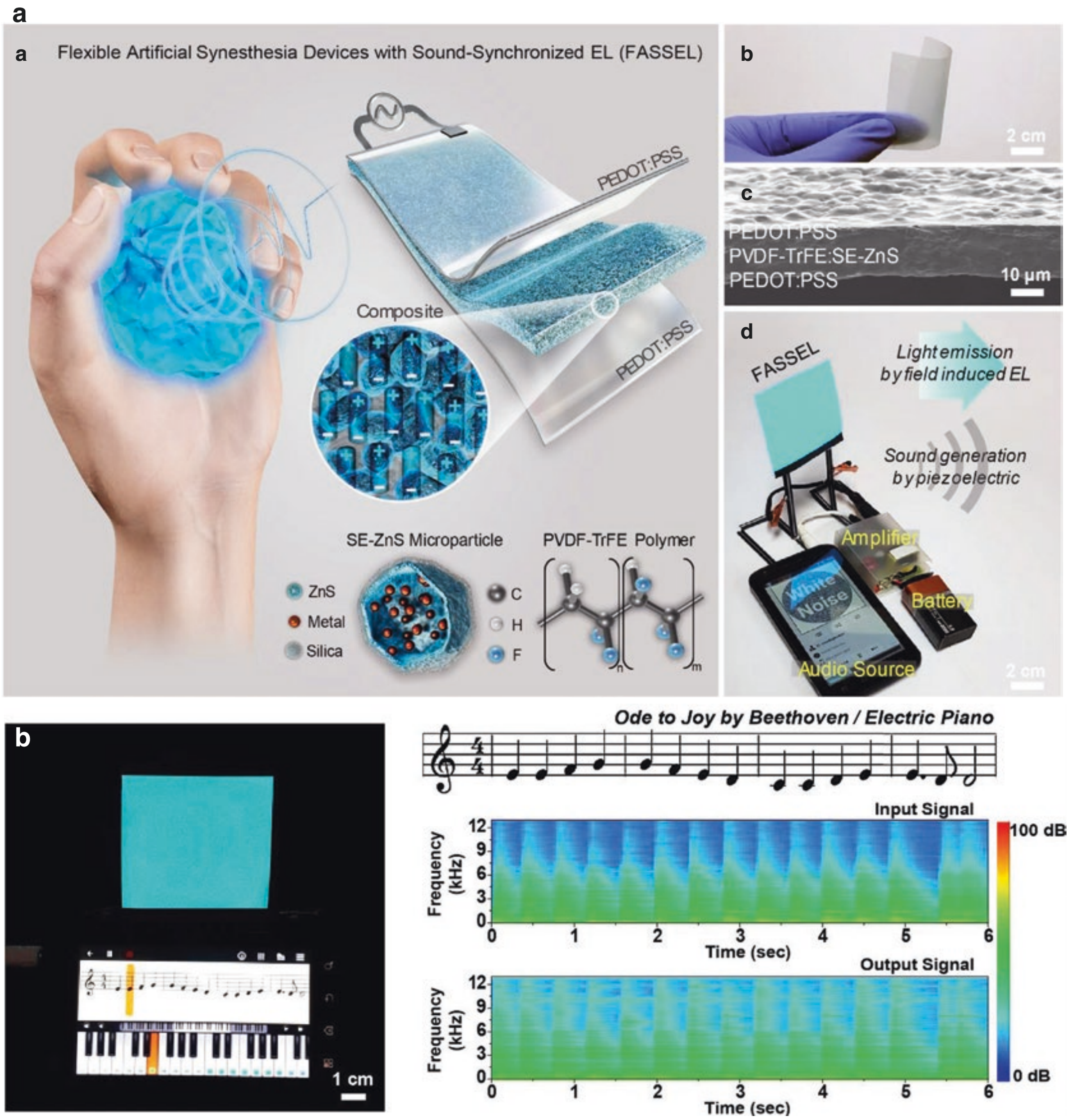

Fig. 7.5 (**a**) Thin piezoelectric film-based flexible artificial synesthesia device with sound-synchronized electroluminescence (FASSEL) (From Kim et al. (2019), Nano Energy, with permission). (**b**) Demonstration of music "Ode to Joy" by the FASSEL. (From Kim et al. (2019), Nano Energy, with permission)

working mechanism is as follows: visual luminance of FASSEL changes as a function of sound frequency, and the color-alteration is frequency-dependent. As a demonstration, while playing the music of “Ode to Joy,” the FASSEL exhibited various colors with different brightness (Fig. 7.5b).

7.9 Summary: New Directions for Synesthetic Correspondences

7.9.1 Synesthetic Correspondence in Embodied Cognition and Human-Machine (AI) Interaction

The synesthetic correspondence so far has been mainly materialized in perceptual domain in which multiple sensory events/stimuli interact. The sensory stimuli in interest are less relevant with humans’ active actions. However, our brain could distinguish between self- and other—generated sensory signals which lead to effects such as sensory attenuation (such as one cannot tickle himself/herself) or intentional binding (temporal interval between one’s own voluntary action and its sensory consequences is subjectively compressed) (Zeng and Chen 2019). There remains the possibility that one’s own action could correspond to the sensory consequences from his/her own action, or from others (Legaspi et al. 2019). Indeed, observers’ uncertainty about their reported perceptual estimate reflects their perceived causal uncertainty. The framework of embodied or grounded cognition suggests that creative thoughts (such as improvising a piano solo) are partially served by simulations of motor activity associated with tools and their use (Matheson and Kenett 2020). Some types of neurodegenerative disorders, including Parkinson disorder can be interpreted within an embodied cognition framework (Bocanegra et al. 2015). At its core, higher order cognitive functions are grounded in the sensory-motor system (Gallese and Cuccio 2018). Therefore, a potential new form of correspondence between action semantics and sensory outputs (representations) within the sensorimotor loop, though uncharted so far, will surely help to address more fundamental question in consciousness and even neuroscience as a whole.

7.9.2 Knowledge (Concepts) Representation

The correspondence between sensory properties across different levels suggests that our concepts or knowledge (representation) of senses need to be enriched and updated. The representation of each sensory event (category) may not be independent and the knowledge of sensory properties could be well understood in terms of the “interaction” of semantic vocabularies. With the corpus exploration method, Lynott and Connell (2009) provided a set of norms for 423 adjectives, each describing an object property. They mapped the subjective sensory ratings (experiences) of the “word,” according to the five perceptual modalities (visual, haptic, auditory, olfactory, and gustatory). Likewise, the data set also included the “exclusivity” of sensory modalities—that one adjective could be only measured/experienced through unisensory perception rather than cross-modally (Lynott and Connell 2009).

This approach of corpus lexicon exploration gives a hint to instruct how well properties from two sensory modalities are corresponded.

In the following study, Lynott and Connell (2013) further revealed that noun concepts are more multimodal than adjective concepts, as nouns tend to subsume multiple adjective property concepts. For example, as a noun, “baby” involves auditory, haptic, olfactory, and visual properties or elicits imagination of those properties—leading to multimodal perceptual strength

(Lynott and Connell 2013). By discovering modality-specific norms of perceptual strength, it is useful for exploring not just the nature of grounded concepts, but also the nature of *form–meaning* relationships, which should further determine the correspondence between two identities (nouns).

Iatropoulos et al. (2018) developed a novel computational method to characterize the olfaction-related semantic content of words. They used a large text corpus of internet sites in English. Two new metrics: olfactory association index (OAI) and olfactory specificity index (OSI) were employed. OAI is targeting on the "strength" that a word is associated with olfaction, while OSI is measuring the specificity of describing given odors. Iatropoulos et al. (2018) further validated the utilities of using OAI and OSI by showing the higher correlations between seemingly scores of given terms and subjective reports of odor association/familiarity. Therefore, this piece of study suggested that OAI and OSI could be good tools to investigate olfactory perception and cognition, including setting up guideline for correspondences between (semantic) odors and even across different sensory stimuli (vocabularies) (Iatropoulos et al. 2018).

Most people find it profoundly difficult to name familiar smells or tactile perception. This difficulty persists even when they received the test in nearly undisturbed conditions, with normal performance of naming visual objects. The essential scientific question is that through what stages and neural (language) pathways that the correspondence is achieved? A recent neuroimaging study suggested they may contain three processes: the object perception, lexical-semantic integration, and verbalization (Olofsson and Gottfried 2015). This line of framework also helps to elucidate the naming difficulty in other sensory modalities (including tactile modality) and find potential solutions. Broadly speaking, human sensory experience also shapes what and how knowledge is stored/represented in our brain. Wang et al. (2020) have proposed sensory-derived and language- and cognition-derived knowledge representations in the human brain: with different underlying brain systems (Wang et al. 2020). Specifically, they recruited congenitally (or early) blind individuals and normal sighted individuals and compared the brain basis of color knowledge in these two groups. For the congenitally blind individuals, their color knowledge can only be obtained through language descriptions and/or cognitive inference, while for the normal peer, their knowledge benefits from both sensory experience and language. Wang et al. (2020) revealed evidence of a sensory-independent knowledge coding system in both groups. Therefore, a comprehensive approach to integrate knowledge graph and artificial intelligence, with neuroscience methodology (neuroimaging), will help to establish the principle of knowledge representations in the human brain, and decipher the intricate relationship between knowledge representation and (synesthetic) sensory correspondence in brain systems.

7.9.3 Synesthetic Correspondence, Metacognition, and Multisensory Integration

As implicated above, our knowledge about properties (relationship) among different sensory stimuli may be abstract. The underlying cognitive ability of knowledge (inference) could be included as one facet of metacognition.

Metacognition—knowing about the ability to monitor one's own decisions and representations, their accuracy and uncertainty—is considered a hallmark of intelligent behavior. The crossmodal correspondences between different sensory attributes, usually relate to one's inferring of the world's causal structure. On the one hand, the underlying perceptual inference is detailed and concrete though with uncertainty and sometime ambiguous in "one-to-one" mapping between sensory properties from different domains (Deroy et al. 2016a, b). On the other hand, the multiple sorts of intersensory correspondences (Spence and Deroy 2013), such as spatiotemporal coincidence, semantic and other higher order correspondences, can inform/teach the brain about whether signals are likely to stem from a common source or independent ones (with the possibility of "averaging" different sensory cues). Metacognition hence provides a constructive tool to integrate information from various sensory channels, supported by the framework of causal inference and cue combination. Specifically, the causal metacognition emphasizes the fundamental links between elementary perception and higher level cognitive processing including social and abstract reasoning. In the framework of Bayesian inference, metacognition has been deciphered with three components that could gain access to monitoring human estimates as well as associated uncertainties: forced-fusion and full-segregation spatial estimates; the inferred causal structure; and the final cross-modal (spatial/temporal) estimates with Bayesian causal inference (Deroy et al. 2016a, b).

Causal metacognition hence is very promising for our understanding of a higher, common level of multisensory processing and could accommodate well the atypical developmental related sensory processing such as in schizophrenia, who had malfunctions in multisensory binding, causal inference, and metacognitive control. That said, to the best of our knowledge, no report of computational modeling for synesthetic correspondence with metacognition component has been given.

7.9.4 Cross-modal Correspondence in Virtual Reality

For the present, most cross-modal (synesthetic) correspondence has been implemented in 2-D scenarios even with higher (semantic) constructs. To extend the exploration in 3-D environment, Huang et al. (2020) implemented a virtual reality (VR) study to examine the color-flavor associations. In this study, participants were not given any visual cues of the green or red tea during drinking but were encouraged to report the color that immediately came to their mind. There remained relatively stable correspondence: green tea was more greenish than red tea, suggesting this correspondence could be contributed by the relevant properties of the bitterness and astringency along with the color dimension of green/red (Huang et al. 2020). Sensory correspondence can be effectively investigated in virtual environment and largely resembling the actual life conditions, indicating a feasible direction of sensory branding (with necessary sensory properties modulation in VR) research and practice.

7.10 Concluding Remarks

Synesthetic correspondence in cross-modal processing represents a vibrant topic in multisensory studies. The "correspondence" concept was originally proposed and henceforth extensively exploited in visual domain (typically in visual apparent motion). It has now been extended and points to the mapping of all discovered sensory properties, explicitly or implicitly across different sensory domains. The inherent "correspondence" contains at least three levels, from concrete levels to abstract ones: structural, statistical, and semantic. With the burgeon of artificial intelligence technology (including virtual reality) in human-machine interaction, as well as the our deepened endogenous understanding—such as metacognition (Deroy et al. 2016a, b) and hierarchical Bayesian inference (Rohe et al. 2019) in multisensory perception, the studies of synesthetic correspondence and potential applications will surely burgeon, and help us to discover the hidden knowledge of sensory representations, as well as to enrich the vocabulary for deciphering the intricate relationship between sensory events during multisensory life.

Acknowledgments This work was supported by the STI2030-Major Project 2021ZD0202600.

References

Agapakis CM, Tolaas S (2012) Smelling in multiple dimensions. Curr Opin Chem Biol 16(5–6):569–575. https://doi.org/10.1016/j.cbpa.2012.10.035

Andrea G (2007) Color names and dynamic imagery. In speaking of colors and odors. In: Plümacher M, Holz P (eds) Converging evidence in language and communication research, vol 8. John Benjamins, Amsterdam, pp 129–140

Barton DN, Halberstadt J (2018) A social Bouba/Kiki effect: a bias for people whose names match their faces. Psychon Bull Rev 25(3):1013–1020. https://doi.org/10.3758/s13423-017-1304-x

Bertelson P, Aschersleben G (2003) Temporal ventriloquism: crossmodal interaction on the time dimension. 1. Evidence from auditory-visual temporal order judgment. Int J Psychophysiol 50(1–2):147–155. https://doi.org/10.1016/s0167-8760(03)00130-2

Bertelson P, Pavani F, Ladavas E, Vroomen J, de Gelder B (2000) Ventriloquism in patients with unilateral visual neglect. Neuropsychologia 38(12):1634–1642. https://doi.org/10.1016/s0028-3932(00)00067-1

Bien N, ten Oever S, Goebel R, Sack AT (2012) The sound of size: crossmodal binding in pitch-size synesthesia: a combined TMS. EEG and psychophysics study Neuroimage 59(1):663–672. https://doi.org/10.1016/j.neuroimage.2011.06.095

Blomert L(2011) The neural signature of orthographic-phonological binding in successful and failing reading development. Neuroimage 57:695–703. https://doi.org/10.1016/j.neuroimage.2010.11.003

Bocanegra, Y., Garcia, A. M., Pineda, D., Buritica, O., Villegas, A., Lopera, F., . . . Ibanez, A. (2015). Syntax, action verbs, action semantics, and object semantics in Parkinson's disease: Dissociability, progression, and executive influences. Cortex, 69, 237–254. https://doi.org/10.1016/j.cortex.2015.05.022

Bottini R, Barilari M, Collignon O (2019) Sound symbolism in sighted and blind. The role of vision and orthography in sound-shape correspondences. Cognition 185:62–70. https://doi.org/10.1016/j.cognition.2019.01.006

Brang D, Ghiam M, Ramachandran VS (2013) Impaired acquisition of novel grapheme-color correspondences in synesthesia. Front Hum Neurosci 7:717. https://doi.org/10.3389/fnhum.2013.00717

Bremner AJ, Caparos S, Davidoff J, de Fockert J, Linnell KJ, Spence C (2013) "Bouba" and "Kiki" in Namibia? A remote culture make similar shape-sound matches, but different shape-taste matches to westerners. Cognition 126(2):165–172. https://doi.org/10.1016/j.cognition.2012.09.007

Brent B, Paul K (1969) Basic color terms: their universality and evolution. University of California Press, Berkeley

Bueti D, Walsh V (2009) The parietal cortex and the representation of time, space, number and other magnitudes. Philosophical Transactions of the Royal Society B-Biological Sciences 364(1525):1831–1840. https://doi.org/10.1098/rstb.2009.0028

Cassirer E (1953) Language and myth. Dover Publications. Dover ed edition (June 1, 1953)

Chen L, Vroomen J (2013) Intersensory binding across space and time: a tutorial review. Atten Percept Psychophys 75(5):790–811. https://doi.org/10.3758/s13414-013-0475-4

Chen LH, Shi ZH, Muller HJ (2010) Influences of intra-and crossmodal grouping on visual and tactile Ternus apparent motion. Brain Res 1354:152–162. https://doi.org/10.1016/j.brainres.2010.07.064

Chen LH, Zhang ML, Ai F, Xie WY, Meng XZ (2016) Crossmodal synesthetic congruency improves visual timing in dyslexic children. Res Dev Disabil 55:14–26. https://doi.org/10.1016/j.ridd.2016.03.010

Chen LH, Zhou XL, Muller HJ, Shi ZH (2018) What you see depends on what you hear: temporal averaging and Crossmodal integration. Journal of Experimental Psychology-General 147(12):1851–1864. https://doi.org/10.1037/xge0000487

Cuskley C, Simner J, Kirby S (2017) Phonological and orthographic influences in the bouba-kiki effect. Psychol Res 81(1):119–130. https://doi.org/10.1007/s00426-015-0709-2

De Carolis L, Marsico E, Arnaud V, Coupe C (2018) Assessing sound symbolism: investigating phonetic forms, visual shapes and letter fonts in an implicit bouba-kiki experimental paradigm. PLoS One 13(12):e0208874. https://doi.org/10.1371/journal.pone.0208874

Deroy O, Fasiello I, Hayward V, Auvray M (2016a) Differentiated audio-tactile correspondences in sighted and blind individuals. J Exp Psychol Hum Percept Perform 42(8):1204–1214. https://doi.org/10.1037/xhp0000152

Deroy O, Spence C, Noppeney U (2016b) Metacognition in multisensory perception. Trends Cogn Sci 20(10):736–747. https://doi.org/10.1016/j.tics.2016.08.006

Ernst MO, Bülthoff HH (2004) Merging the senses into a robust percept. Trends Cogn Sci 8:162–169. https://doi.org/10.1016/j.tics.2004.02.002

Favalli S, Skov T, Spence C, Byrne DV (2013) Do you say it like you eat it? The sound symbolism of food names and its role in the multisensory product experience. Food Res Int 54(1):760–771. https://doi.org/10.1016/j.foodres.2013.08.022

Fort M, Martin A, Peperkamp S (2015) Consonants are more important than vowels in the Bouba-kiki effect. Lang Speech 58(Pt 2):247–266. https://doi.org/10.1177/0023830914534951

Fryer L, Freeman J, Pring L (2014) Touching words is not enough: how visual experience influences haptic-auditory associations in the "Bouba-Kiki" effect. Cognition 132(2):164–173. https://doi.org/10.1016/j.cognition.2014.03.015

Gallese V, Cuccio V (2018) The neural exploitation hypothesis and its implications for an embodied approach to language and cognition: insights from the study of action verbs processing and motor disorders in Parkinson's disease. Cortex 100:215–225. https://doi.org/10.1016/j.cortex.2018.01.010

Gertner L, Henik A, Reznik D, Cohen KR (2013) Implications of number-space synesthesia on the automaticity of numerical processing. Cortex 49:1352–1362. https://doi.org/10.1016/j.cortex.2012.03.019

Gold R, Segal O (2017) The bouba-kiki effect and its relation to the autism quotient (AQ) in autistic adolescents. Res Dev Disabil 71:11–17. https://doi.org/10.1016/j.ridd.2017.09.017

Goodhew SC, Freire MR, Edwards M (2015) Enhanced semantic priming in synesthetes independent of sensory binding. Conscious Cogn 33:443–456. https://doi.org/10.1016/j.concog.2015.02.019

Grossenbacher PG, Lovelace CT (2001) Mechanisms of synesthesia: cognitive and physiological constraints. Trends Cogn Sci 5(1):36–41. https://doi.org/10.1016/s1364-6613(00)01571-0

Guo L, Bao M, Guan LY, Chen LH (2017) Cognitive styles differentiate crossmodal correspondences between pitch glide and visual apparent motion. Multisens Res 30(3–5):363–385

Haryu E, Kajikawa S (2012) Are higher-frequency sounds brighter in color and smaller in size? Auditory-visual correspondences in 10-month-old infants. Infant Behavior & Development 35:727–732. https://doi.org/10.1016/j.infbeh.2012.07.015

Hein E, Cavanagh P (2012) Motion correspondence in the Ternus display shows feature bias in spatiotopic coordinates. J Vis 12(7):16. https://doi.org/10.1167/12.7.16

Hidaka S, Yaguchi A (2018) An investigation of the relationships between autistic traits and crossmodal correspondences in typically developing adults. Multisens Res 31(8):729–751. https://doi.org/10.1163/22134808-20181304

Huang FX, Qi YX, Wang CJ, Wan XA (2020) Show me the color in your mind: a study of color-flavor associations in virtual reality. Food Qual Prefer 85:103969. https://doi.org/10.1016/j.foodqual.2020.103969

Hubbard EM, Arman AC, Ramachandran VS, Boynton GM (2005) Individual differences among grapheme-color synesthetes: brain-behavior correlations. Neuron 45(6):975–985

Hubbard EM, Brang D, Ramachandran VS (2011) The cross-activation theory at 10. J Neuropsychol 5:152–177

Hung SM, Styles SJ, Hsieh PJ (2017) Can a word sound like a shape before you have seen it? Sound-shape mapping prior to conscious awareness. Psychol Sci 28(3):263–275. https://doi.org/10.1177/0956797616677313

Iatropoulos G, Herman P, Lansner A, Karlgren J, Larsson M, Olofsson JK (2018) The language of smell: connecting linguistic and psychophysical properties of odor descriptors. Cognition 178:37–49. https://doi.org/10.1016/j.cognition.2018.05.007

Judd DB, Wyszecki G (1975) Color in business, science, and industry. Wiley, New York

Karthikeyan S, Rammairone B, Ramachandra V (2016) The Bouba-Kiki phenomenon tested via schematic drawings of facial expressions: further validation of the internal simulation hypothesis. Iperception 7(1):2041669516631877. https://doi.org/10.1177/2041669516631877

Ketron S, Spears N (2019) Sounds like a heuristic! Investigating the effect of sound-symbolic correspondences between store names and sizes on consumer willingness-to-pay. J Retail Consum Serv 51:285–292. https://doi.org/10.1016/j.jretconser.2019.06.016

Kim, J. S., Cho, S. H., Kim, K. L., Kim, G., Lee, S. W., Kim, E. H., . . . Park, C. (2019). Flexible artificial synesthesia electronics with sound-synchronized electroluminescence. Nano Energy, 59, 773–783. doi:https://doi.org/10.1016/j.nanoen.2019.03.006

Köhler W (1929) Gestalt psychology. Liveright, New York

Kohler (1947) Gestalt psychology, 2nd edn. Liveright Publishing, New York, NY

Legaspi R, He ZQ, Toyoizumi T (2019) Synthetic agency: sense of agency in artificial intelligence. Curr Opin Behav Sci 29:84–90. https://doi.org/10.1016/j.cobeha.2019.04.004

Levinson SC, Majid A (2014) Differential ineffability and the senses. Mind Lang 29(4):407–427. https://doi.org/10.1111/mila.12057

Lynott D, Connell L (2009) Modality exclusivity norms for 423 object properties. Behav Res Methods 41(2):558–564. https://doi.org/10.3758/Brm.41.2.558

Lynott D, Connell L (2013) Modality exclusivity norms for 400 nouns: the relationship between perceptual experience and surface word form. Behav Res Methods 45(2):516–526. https://doi.org/10.3758/s13428-012-0267-0

Macdermott MM (1940) Vowel sounds in poetry: their music and tone-colour. Kegan Paul, London

Marks LE (2014) The Unity of the senses: interrelations among the modalities. Academic Press

Marks L, Mulvenna C (2013) Synesthesia, at and near its borders. Front Psychol 4:651. https://doi.org/10.3389/fpsyg.2013.00651

Martino G, Marks LE (1999) Perceptual and linguistic interactions in speeded classification: tests of the semantic coding hypothesis. Perception 28(7):903–923. https://doi.org/10.1068/p2866

Matheson HE, Kenett YN (2020) The role of the motor system in generating creative thoughts. NeuroImage 213:116697. https://doi.org/10.1016/j.neuroimage.2020.116697

Maurer D, Pathman T, Mondloch CJ (2006) The shape of boubas: sound-shape correspondences in toddlers and adults. Dev Sci 9(3):316–322. https://doi.org/10.1111/j.1467-7687.2006.00495.x

McCarthy JD, Caplovitz GP (2014) Color synesthesia improves color but impairs motion perception. Trends Cogn Sci 18(5):224–226. https://doi.org/10.1016/j.tics.2014.02.002

Morein-Zamir S, Soto-Faraco S, Kingstone A (2003) Auditory capture of vision: examining temporal ventriloquism. Brain Res Cogn Brain Res 17(1):154–163. https://doi.org/10.1016/s0926-6410(03)00089-2

Ngo MK, Velasco C, Salgado A, Boehm E, O'Neill D, Spence C (2013) Assessing crossmodal correspondences in exotic fruit juices: the case of shape and sound symbolism. Food Qual Prefer 28(1):361–369. https://doi.org/10.1016/j.foodqual.2012.10.004

Occelli V, Esposito G, Venuti P, Walker P, Zampini M (2012). Audiovisual crossmodal correspondences in Autism Spectrum Disorders (ASDs). Seeing and Perceiving, 25:44–44 https://doi.org/10.1163/187847612X646668

Ogulmus C, Karacaoglu M, Kafaligonul H (2018) Temporal ventriloquism along the path of appar-

ent motion: speed perception under different spatial grouping principles. Exp Brain Res 236(3):629–643. https://doi.org/10.1007/s00221-017-5159-1
Olofsson JK, Gottfried JA (2015) The muted sense: neurocognitive limitations of olfactory language. Trends Cogn Sci 19(6):314–321. https://doi.org/10.1016/j.tics.2015.04.007
Orchard-Mills E, Van der Burg E, Alais D (2016) Crossmodal correspondence between auditory pitch and visual elevation affects temporal ventriloquism. Perception 45(4):409–424. https://doi.org/10.1177/0301006615622320
Palmer SE, Schloss KB, Xu Z, Prado-Leon LR (2013) Music-color associations are mediated by emotion. Proc Natl Acad Sci U S A 110(22):8836–8841. https://doi.org/10.1073/pnas.1212562110
Parise CV (2016) Crossmodal correspondences: standing issues and experimental guidelines. Multisens Res 29(1–3):7–28. https://doi.org/10.1163/22134808-00002502
Parise CV, Ernst MO (2016) Correlation detection as a general mechanism for multisensory integration. Nat Commun 7:11543. https://doi.org/10.1038/ncomms11543
Parise C, Spence C (2008) Synesthetic congruency modulates the temporal ventriloquism effect. Neurosci Lett 442(3):257–261. https://doi.org/10.1016/j.neulet.2008.07.010
Parise CV, Spence C (2009) 'When birds of a feather flock together': synesthetic correspondences modulate audiovisual integration in non-synesthetes. PLoS One 4(5):e5664. https://doi.org/10.1371/journal.pone.0005664
Parise CV, Spence C (2012) Audiovisual crossmodal correspondences and sound symbolism: a study using the implicit association test. Exp Brain Res 220(3–4):319–333. https://doi.org/10.1007/s00221-012-3140-6
Parise CV, Spence C, Ernst MO (2012) When correlation implies causation in multisensory integration. Curr Biol 22(1):46–49. https://doi.org/10.1016/j.cub.2011.11.039
Peiffer-Smadja N, Cohen L (2019) The cerebral bases of the bouba-kiki effect. NeuroImage 186:679–689. https://doi.org/10.1016/j.neuroimage.2018.11.033
Piqueras-Fiszman B, Spence C (2011) Crossmodal correspondences in product packaging. Assessing color-flavor correspondences for potato chips (crisps). Appetite 57(3):753–757. https://doi.org/10.1016/j.appet.2011.07.012
Rakova M (2003) The extent of the literal: metaphor, polysemy and theories of concepts. Palgrave Macmillan, New York
Ramachandran V, Hubbard EM (2001a) Synaesthesia - a window into perception, thought and language. J Conscious Stud 8(12):3–34
Ramachandran VS, Hubbard EM (2001b) Psychophysical investigations into the neural basis of synaesthesia. Proceedings of the Royal Society B-Biological Sciences 268(1470):979–983
Ramachandran VS, Hubbard EM (2003) Hearing colors, tasting shapes. Sci Am 288(5):52–59. https://doi.org/10.1038/scientificamerican0503-52
Ramachandra V (2016) The linguistic and cognitive factors associated with lexical-gustatory synesthesia: A case study. Brain and cognition 106:23–32. https://doi.org/10.1016/j.bandc.2016.04.005
Roblee L, Washburn MF (1912) Minor studies from the psychological laboratory of Vassar college XX the affective values of articulate sounds. Am J Psychol 23:579–583. https://doi.org/10.2307/1413063
Rohe T, Ehlis AC, Noppeney U (2019) The neural dynamics of hierarchical Bayesian causal inference in multisensory perception. Nat Commun 10(1):1907. https://doi.org/10.1038/s41467-019-09664-2
Ross C (2018) Unlocking consciousness- lessons from the convergence of computing and cognitive psychology. World Scientific
Rothen N, Schwartzman DJ, Bor D, Seth AK (2018) Coordinated neural, behavioral, and phenomenological changes in perceptual plasticity through overtraining of synesthetic associations. Neuropsychologia 111:151–162. https://doi.org/10.1016/j.neuropsychologia.2018.01.030
Rouw R, van Driel J, Knip K, Richard RiK (2013) Executive functions in synesthesia. Consciousness and cognition 22:184–202. https://doi.org/10.1016/j.concog.2012.11.008
Rouw R, Scholte HS (2016) Personality and cognitive profiles of a general synesthetic trait. Neuropsychologia 88:35–48. https://doi.org/10.1016/j.neuropsychologia.2016.01.006
Sadaghiani S, Maier JX, Noppeney U (2009) Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. J Neurosci 29(20):6490–6499. https://doi.org/10.1523/Jneurosci.5437-08.2009
Saenz M, Koch C (2008) The sound of change: visually-induced auditory synesthesia. Curr Biol 18(15):R650–R651. https://doi.org/10.1016/j.cub.2008.06.014
Sakamoto M, Watanabe J (2018) Bouba/Kiki in touch: associations between tactile perceptual qualities and Japanese phonemes. Front Psychol 9:295. https://doi.org/10.3389/fpsyg.2018.00295
Shen Y (2008) Metaphor and poetic figures. The Cambridge handbook of metaphor and thought. Cambridge University Press. 295–307
Siegfried W (2007) Color terms between elegance and beauty: the verbalization of color with textiles and cosmetics. In: Plümacher M, Holz P (eds) Speaking of colors and odors. [converging evidence in language and communication research 8]. John Benjamins, Amsterdam, pp 113–128
Slutsky DA, Recanzone GH (2001) Temporal and spatial dependency of the ventriloquism effect. Neuroreport 12(1):7–10. https://doi.org/10.1097/00001756-200101220-00009
Spence C (2011) Crossmodal correspondences: a tutorial review. Atten Percept Psychophys 73(4):971–995. https://doi.org/10.3758/s13414-010-0073-7

Spence C (2012) Managing sensory expectations concerning products and brands: capitalizing on the potential of sound and shape symbolism. J Consum Psychol 22(1):37–54. https://doi.org/10.1016/j.jcps.2011.09.004
Spence C, Deroy O (2013) How automatic are crossmodal correspondences? Conscious Cogn 22(1):245–260. https://doi.org/10.1016/j.concog.2012.12.006
Staniewski P (2017) Geschmack und smak– sprachliche Aspekte der gustatorischen Wahrnehmung im Deutschen und Polnischen. In: Szczek J, Kalasznik M (eds) Beitrage zur Fremdsprachenvermittlung. Verlag Empirische Pädagogik, Landau, pp 223–248
Stein BE, Meredith MA (1993) The merging of the senses. The MIT Press, Cambridge, MA
Tian Y, Liu X, Chen L (2020) Mindfulness meditation biases visual temporal order discrimination but not under conditions of temporal ventriloquism. Front Psychol 11:1937. https://doi.org/10.3389/fpsyg.2020.01937
Tzeng CY, Nygaard LC, Namy LL (2017) The Specificity of Sound Symbolic Correspondences in Spoken Language. Cognitive science 41:2191–2220. https://doi.org/10.1111/cogs.12474
Ullman S (1979) The interpretation of visual motion. MIT Press, Cambridge, MA
van Petersen E, Altgassen M, van Lier R, van Leeuwen TM (2020) Enhanced spatial navigation skills in sequence-space synesthetes. Cortex 130:49–63. https://doi.org/10.1016/j.cortex.2020.04.034
Vroomen J, Keetels M (2006) The spatial constraint in intersensory pairing: no role in temporal ventriloquism. J Exp Psychol Hum Percept Perform 32(4):1063–1071. https://doi.org/10.1037/0096-1523.32.4.1063
Vroomen J, de Gelder B, Vroomen J (2004) Temporal ventriloquism: sound modulates the flash-lag effect. J Exp Psychol Hum Percept Perform 30(3):513–518. https://doi.org/10.1037/0096-1523.30.3.513
Walsh V (2003) A theory of magnitude: common cortical metrics of time, space and quantity. Trends Cogn Sci 7(11):483–488. https://doi.org/10.1016/j.tics.2003.09.002
Wan X, Woods AT, van den Bosch JJ, McKenzie KJ, Velasco C, Spence C (2014) Cross-cultural differences in crossmodal correspondences between basic tastes and visual features. Front Psychol 5:1365. https://doi.org/10.3389/fpsyg.2014.01365
Wang QC, Bao M, Chen LH (2014) The role of spatio-temporal and spectral cues in segregating short sound events: evidence from auditory Ternus display. Exp Brain Res 232(1):273–282. https://doi.org/10.1007/s00221-013-3738-3
Wang XY, Men WW, Gao JH, Caramazza A, Bi YC (2020) Two forms of knowledge representations in the human brain. Neuron 107(2):383. https://doi.org/10.1016/j.neuron.2020.04.010
Westbury C (2005) Implicit sound symbolism in lexical access: evidence from an interference task. Brain Lang 93(1):10–19. https://doi.org/10.1016/j.bandl.2004.07.006
Westbury C, Hollis G, Sidhu DM, Pexman PM (2018) Weighing up the evidence for sound symbolism: distributional properties predict cue strength. J Mem Lang 99:122–150
Winter B, Marghetis T, Matlock T (2015) Of magnitudes and metaphors: explaining cognitive interactions between space, time, and number. Cortex 64:209–224. https://doi.org/10.1016/j.cortex.2014.10.015
Yeshurun Y, Sobel N (2010) An odor is not worth a thousand words: from multidimensional odors to unidimensional odor objects. Annu Rev Psychol 61(219–241):C211–C215. https://doi.org/10.1146/annurev.psych.60.110707.163639
Zeljko M, Kritikos A, Grove PM (2019) Lightness/pitch and elevation/pitch crossmodal correspondences are low-level sensory effects. Atten Percept Psychophys 81(5):1609–1623. https://doi.org/10.3758/s13414-019-01668-w
Zeng HK, Chen LH (2019) Robust temporal averaging of time intervals between action and sensation. Front Psychol 10:511. https://doi.org/10.3389/fpsyg.2019.00511
Zhang Y, Chen L (2016) Crossmodal statistical binding of temporal information and stimuli properties recalibrates perception of visual apparent motion. Front Psychol 7:434. https://doi.org/10.3389/fpsyg.2016.00434
Zhao QQ, Huang CR, Ahrens K (2019) Directionality of linguistic synesthesia in mandarin: a corpus-based study. Lingua 232:102744. https://doi.org/10.1016/j.lingua.2019.102744