Beyond Vision: Extending the Scope of a Sensorimotor Account of Perception

Caroline Lyon

Adaptive Systems Research Group, University of Hertfordshire, Hatfield, UK
C.Lyon@herts.ac.uk

Abstract. We examine the scope of some sensorimotor accounts of perception, and their application in developmental robotics. Current interest in sensorimotor theories, and the enactive paradigm, was stimulated by the seminal book *The Embodied Mind* by Varela, Thompson and Rosch (1991) [32]. However, both in this initial book and subsequently there has been much work on visual perception and less attention to other perceptual modalities. We suggest that the insights gained from an exploration of the visual domain need supplementing, and in some respects qualifying: some significant characteristics of vision do not hold for audition, in particular for the perception of speech. This leads into a discussion of the importance of integrating different perceptual modes, with particular reference to robots and human-robot interaction. We examine the effect of including audition in accounts of perception, and suggest that it makes sense to avoid the unnecessary straight jacket of a model based primarily on vision and touch alone. The sensorimotor approach can be extended to other perceptual modes.

Keywords: sensorimotor, perception, vision, audition, human-robot interaction.

1 Introduction

In this chapter we examine the scope of some sensorimotor accounts of perception, and their application in developmental robotics. The current interest in sensorimotor theories, and the enactive paradigm, was stimulated by the seminal book *The Embodied Mind* by Varela, Thompson and Rosch (1991) [32]. However, both in this initial book and subsequently there has been much work on visual perception (for example [25,24]), some on touch, but little attention to other perceptual modalities. Varela et al. write of perception in general, but focus on vision and explore in detail the perception of color. They propose that

> color provides a paradigm of a cognitive domain that is neither pregiven nor represented but rather experiential and enacted ..... The time has come, however, to step back and consider some of the lessons this cognitive domain provides for our understanding of perception and cognition in general. [32, page 171]

We suggest that the insights gained from an exploration of the visual domain need supplementing, and in some respects qualifying: when we examine auditory perception we find that some significant characteristics of vision do not hold for audition.

J.M. Bishop and A.O. Martin (eds.), *Contemporary Sensorimotor Theory,*
Studies in Applied Philosophy, Epistemology and Rational Ethics 15,
DOI: 10.1007/978-3-319-05107-9_9,

This leads into a discussion of the importance of integrating different perceptual modes, with particular reference to human-robot interaction. Starting from the philosophical origins of sensorimotor theory we pick up some of the ideas that turn out to be relevant to present day issues. We examine the effect of including audio and other perceptual modes in accounts of perception, and suggest that it makes sense to avoid the unnecessary straight jacket of a model based primarily vision and touch. The sensorimotor approach can be extended to other perceptual modes.

2 Background

The genesis of ideas expounded by Varela et al., and their followers, can be traced back through European philosophers Merleau-Ponty, Heidegger, to Husserl [19,9,12]. (For an accesible overview see [18].) They were also influenced by strands of Buddhist thought. The interest in vision was characteristic of all these philosophers, and a similar trend was evident in the empiricist British school, typified by works such as those of Berkeley [4], Locke [16] and, to some extent, Hume [11]. For instance, Berkeley produced his *New Theory of Vision*, where "new" was 1709. Locke wrote of "sight, the most comprehensive of all our senses ...".

Thus Varela, Thompson and Rosch continued in a field which had given a pre-eminent position to vision. They set out to counter cognitivist approaches that were influential in the latter 20th century - representational theories that typically proposed some inner picture or symbol mediating between the outside world and the mind. Their insights into the enactive, embodied nature of perception entailed a different account of visual perception.

Varela et al. acknowledge their debt to Merleau-Ponty, who developed his ideas through critiques of the phenomenology of Husserl and Heidegger [19,20]. Merleau-Ponty explained perception as an embodied activity through which we relate to things in the world around us. As he says in *Phenomenology of Perception* "Perception opens a window onto things. This means it is directed, quasi-teleologically, towards *truth in itself* in which the reason underlying all appearances is to be found" ([19], his italics). Though he goes on to talk about perception in general terms "a window onto things" imply vision and touch, not including hearing, tasting, smelling.

However, without pursuing the question of what *truth in itself* might mean, we can see how such a philosphy begins to be relevant to the development of artificial cognition in robotics. There is no homunculus or inner man viewing percepts that are reconstituted as a model of some part of the environment: the subject is inseparable from the world and "the world is not what I think but what I live through" (ibid, page xviii). This need not only mean a world of visible and tangible "things": different perceptual modalities can, in theory and sometimes in practice, be implemented in a robot and integrated to simulate human cognition [22,33]. Examples can be found in work done in the ITALK project, *Integration and Transfer of Action and Language Knowledge in Robots*, described in [5], in which elements of language are acquired by a humanoid robot interacting with naive human participants, through its own sensorimotor experiences - visual, auditory and proprioceptive. Another example is work done in the SPARK project, described in *Spatial Temporal Patterns for Action-Oriented*

Perception in Roving Robots [1] in which the “agent transforms sensory signals to give rise to motor output ... there is no need for an internal model. Perception is active” (ibid, page viii).

The Focus on Vision

Merleau-Ponty provides the starting point for Varela’s philosophy and subsequent developments in sensorimotor theories, which have become focused predominantly on visual perception. In “The Embodied Mind” [32] Varela et al. take color as a case study. Their illuminating account reviews many experimental results showing how the perception of color is a perceived attribute, partially dependent on the observer and on ambient conditions. The color of an object is seen as part of “a patchwork of visual modalities” including size, shape, motion lighting conditions, etc. (ibid, page 162). Interpretations and associations of color are deeply rooted in our culture, and Varela explores specific cognitive processes related to it[1]. However, though they write of perception in a general sense Varela et al have almost nothing to say about other modes of perception apart from a passing reference to hearing and a paragraph on olfaction. In other work from the Enactivist school the sense of touch is explored, and Noë goes so far as to say that “Touch, not vision, should be our model for perception” [23]. Here we again have perception of “things” that could be seen as well as touched, but excluding audition.

The focus on visual perception is indicated by the titles of writings. For instance, *A sensorimotor account of vision and visual consciousness* by Regan and Noë [25] has been very influential. In the preface to a collection of readings edited by Noë and Thompson entitled *Vision and Mind* [24] the editors say ”The writings in this volume investigate the nature of visual perception. Our goal has been to produce a collection that can serve as a starting point for the philosophy of perception.” We argue that the study of visual perception is indeed a starting point.

However, other modes of perception are often integral parts of the perceptive process, and some of their characteristics, explored below, differ from those of vision. Visual processing is only part of the story.

3 The Need to Integrate Multiple Modes of Perception

Though vision is a key mode of perception, in humans and other animals it is critically integrated with hearing, touching, tasting, smelling as well as with internal proprioceptive information. Vision can often not be disassociated from other perceptual modes and the need to integrate them has underscored much recent work in robotics. For example, in work on language acquisition through interaction between humans and a humanoid robot an acoustic sound stream, visual percepts and proprioceptive information have to be integrated [30]. See Figure 1. Vision plays a significant role in language acquisition

[1] A striking anecdote related to the author concerns a dictat during the Cultural Revolution in China. At that time red was the color of revolution and progress, so it was deemed incorrect to have red mean “stop” on traffic lights. A decree went out that red should mean “go” and green should mean “stop”. Chaos ensued until even the most committed revolutionary agreed to reverse the order.

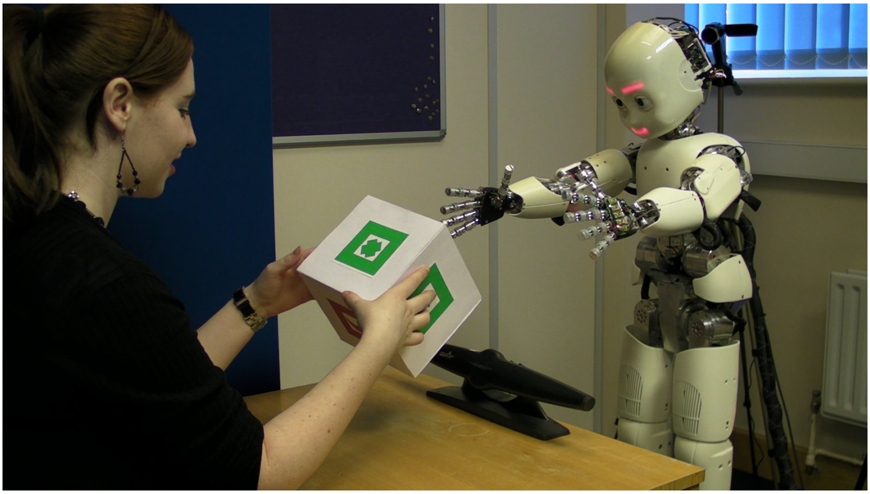

Fig. 1. Experiment with the iCub robot in which a participant teaches it the names of shapes and colors, using visual, auditory and proprioceptive modes of perception

(though not an essential one, as people born blind can learn to speak). For example, research has shown how infants are aware of the gaze of a carer, and shared gaze contributes to language learning.

There are also examples of the integration of different perceptual modes for non-speech sounds. Experiments have been carried out to investigate *vision - action - sound* brain functions using fMRI technology [26]. In observing a familiar action that produces a sound, such as drumming, we can usually predict the sound, but when the natural synchronization was disrupted this had a measurable effect. A second set of experiments investigated the effect on brain processes of disrupting the natural covariance between the velocity of the drumming action and the sound intensity. In both cases expert drummers were compared to non-musicians, and the effects were found to differ. A significant finding was that in synchronised drumming the brain activity in experts was relatively reduced. It was suggested that "the reduced activity found in certain brain areas of musicians was accompanied by an increase of activity in other areas" (ibid, page 1490).

Other examples of the need to study perceptual integration are ubiquitous in robotics. For instance, in modelling a grasping action hand-eye co-ordination has to be mastered [31]. The biological inspiration behind some robotic research has looked to non-human models and in exploring perceptual machines inspired by insects many multimodal interactions between different sensory systems have to be studied: visual, auditory, olfactory, mechanosensory [1].

It is also interesting to consider neuromotor prosthesis studies, such as work on the control of a robotic limb replacing an amputated arm [8]. This is a prime example of a sensorimotor system, based on the capture of neural codes in the subject, and the creation of links between these neural signals and action in the outside world. With closed-loop control the neural activity of the subject guides a device and receives sensory feedback. The perception of this sensory feedback in turn determines the next step in an interactive cycle controlling the movement of the robotic limb. Perception in this case is highly directed, intentional, and active, integrating visual, haptic and proprioceptive percepts.

In a critique of human manual activity Hutto questions whether intentional instructions can control hand movements [13]. “Only very fine-grained instructions would be capable of directing or controlling specific acts of manual activity successfully. This raises a number of questions. How do brains decide which general kind of motor act, M, is the appropriate sort of motor act to use in the situation at hand?”. The control of robotic limbs provides an existential answer to this question: it is through the multisensory feedback cycle.

Tracing back to the ideas of Husserl and Heidigger we can see a link from some of the ideas they examined to issues for neuromotor prosthesis. Grasping actions that are normally done without thinking have to be executed with directed concentration, which is crucial for successful execution. The aim is to move beyond this stage so that actions at a low level become routine, subconscious processes. Heidigger identified these different modes, comparing the routine, subconscious use of a hammer by an experienced carpenter with conscious, directed attention.

Husserl emphasized “intentionality” as a key component of perception. As Dreyfus says, Husserl made intentionality one of the main topics of philosophy [18, page 256]. Now, it is not uncommon in this field to find confusion generated by ambiguous words, and we need to examine the word “intentional”. As well as meaning “directed attention” this also can mean doing something planned [32, page 16], and it is this second meaning which has wider currency in the world at large; for instance, common legal definitions of crimes may include the requirement that they are “intentional” acts. This second meaning is also a crucial concept in neuroscientific modelling (for example in speech production, e.g. [10]) and in robotic development (for example trajectory planning for grasping e.g. [31]) where goal-oriented, forward plans are required. We need to be aware whether “intentional” is used in a present temporal frame as “directed” or in a future temporal frame as “planned”.

Active Visual Processing and Passive Perception in Other Modalities

A key characteristic of visual processing is that it is typically an *active* process. As Varela puts it: ”[t]he enactive approach underscores the importance of two interrelated points: 1) perception consists of perceptually guided action and 2) cognitive structures emerge from the recurrent sensorimotor patterns that enable action to be perceptually guided [32, page 173]. Noë summarises the position succintly as he says: “Perception is not something that happens to us, or in us. It is something we do.” [23]. Perception is a kind of thoughtful activity.

Passive perception can occur occasionally in vision: for instance, if an unexpected flash of light occurs. Passive perception can be exploited in subliminal advertising. However these examples are rare, and typical visual processing is active as articulated by Varela, Noë and others.

Now, contrast this with auditory perception. In this case the hearer can often have a passive as well as an active role. Active listening occurs when a subject directs their attention to speech or other sounds, so humanoid robots acquiring linguistic skills through interaction with a human will typically be listening in an active mode [17]. On the other hand, auditory perception can be passive: a sleeping person can be woken by sounds, and some loud or incessant noise can be an unpleasant experience that the hearer cannot escape. Similarly, perceptions of touch can be active, as in feeling an object, but also passive as in receiving a blow. Perceptions of smell can be active, as in sniffing out the source of an odour, but also passive as when a smell unexpectedly hits your senses.

A striking example of passive perception comes from experiments on the perception of speech sounds as reflected by mismatch negativity (MMN) signals, a component of auditory event-related potential (ERP) [21]. In these experiments human subjects are exposed to a repetitive sequence of standard sounds, interrupted by a different deviant sound. Neuronal populations in the brain encode acoustic invariances specific to each speech sound, and a change in the sequence elicits a measurable signal, the MMN. This occurs whether or not the subject is paying attention. It can even occur in coma patients, in which case it is a predictor of a return to consciousness.

4 The Perception of Phonemes

Consider the perception of phonemes, and the role in language understanding of minimal pairs of words such as *kick* and *lick* or *ball* and *bell*: the change in a single consonant or vowel phoneme alters the meaning. Because of the human ability to distinguish phonemes they were used in the MMN experiments (mentioned above) as well as other sounds.

Interest in phonemes goes back a long way and Sapir's research in the 1930s showed the phoneme is "a cognitive construct that is so strong that it leads individuals to assert the existence of sounds that are not present, and deny the existence of sounds that are present" [29], quoted by [3]. The ability to distinguish between different phonemes is commonly called "categorical perception" in the field of child development, and its acquisition by very young infants is remarkable [28,15,6]. Recent neuroscientific work has identified locations of phoneme detectors in the brain, in both hemispheres [35].

However, it remains the case that phonemes can still only partially be detected by automated recognition systems. In very simple terms, speech recognizers typically have two components: an acoustic processor which extracts a list of candidate phonemes and a language model which compares possible choices of syllables or words (short sequences of phonemes) with stored examples in a very large data base of recorded speech. The output of the recognizer is based on a combination of probabilities derived from these information sources.

The probability of phonemes being correctly recognised by an automated system is related to a number of factors, for instance reading a prepared text gives better results than spontaneous speech. Whether the phoneme is a vowel or a consonant is relevant, as is its position in a syllable. A key factor is the saliency of the word in which the phoneme occurs: content words like nouns, verbs and adjectives are more likely to have canonical, consistent phonemic structure than function words. Research on an analysis of 4 hours of spontaneous telephone speech which was manually annotated by trained phoneticians reported that the word "and" had 80 different phonemic representations [7]. This highlights the point that orthographic transcripts of spontaneous speech may not be a close match to perceived sounds - a point that needs to be taken into account when research into, for instance, child language acquisition uses corpora of written transcripts of child directed speech. The observed fact that content words are more likely to have consistent phonemic structure is exploited in work simulating the learning of word forms by a humanoid robot which interacts with humans [17].

Problems with phoneme recognition are partially due to speakers' pronunciation: human listners use a number of semantic, syntactic and prosodic clues to decode the speech signal. However, it remains the case that at present phonemes cannot reliably be detected by automated recognition systems, even when clearly enunciated. Recent research on phoneme recognition obtains best results of only around 67% [34,2]. This research, based on empirical evidence such as EEG brain wave recordings, suggests that phases of oscillations might account for phonemic discrimination. Port proposes a high-dimensional linguistic memory, incorporating many items of information extracted from an acoustic stream, where phonemes are just statistical invariants drawn from this data [27]. This type of interpretation has led him to question the existence of phonemes, since they are "only" cognitive constructs. There is no reliable external representation of this hearing experience, which contrasts with the traditional sensorimotor view in explaining visual consciousness that "the outside world is its own external representation" [25].

When phonemes are described as *only* cognitive constructs it implies that there could be some other, more real, status. For instance, with a robotic prosthesis we can say it is not real, it contrasts with a real limb. But this is not the case with phonemes: cognitive constructs are as real as you can get. "We must not, therefore, wonder whether we really perceive a world, we must instead say: the world is what we perceive" [32, page xvi].

5 The Varied Meanings of "Representation"

At this point we need to consider some of the varied meanings of the word "representation". It can mean a symbol, as a flag can represent a country. It can mean a pictorial artefact, as in representational art, and an analogous use of the word in this sense was common in cognitivist theories of perception which proposed some sort of mental picture in the brain [32,25,13]. Representations might be thought of as "images, schemas, symbols, models, icons, sentences, maps and so on" [13].

However, the word is also commonly used rather differently to refer to neuronal patterns of activity in the brain. In the report on the drumming research mentioned above the authors write of areas of the brain that "are active in audiomotor, visuomotor and audiovisual *representation* studies" (emphasis added in this and following quotations). Wang [34] describes his work as investigating "the neural *representation* of phonemes". Other examples of the use of the word include: "each sound, both speech and non-speech, develops its neural *representation* corresponding to the percept of this sound" [21]. Another example is: "comprehension is achieved with RH acoustic / phonetic *representations* of speech working in concert with LH mechanisms more sensitve to phonemic category" [35].

In this usage the term “representation” refers to a relationship beween neurons and external stimuli; there is no inner picture or reconstructed model. This contrasts with the use of “representation” to mean some mediating image beween an internal and external world.

With the perception of phonemes there is no identifiable external representation as with visual percepts, but there is a representation in the form of neuronal patterns of activity. Since the advent of brain imaging technologies there has been much research into the neuroscience of language and investigations into neuronal functions. For example, this approach addresses the relationship between speech perception and production, showing how they are critically linked, [10]. Another example reports evidence that phonological input and output buffers hold transient information, and a phonological short term’ memory (pSTM) arises from the cycling of information between the two buffers [14]. Speech processing can be seen as processing statistical invariants extracted from an acoustic signal [27]. But though there is no identifiable external representation we still have sensorimotor interaction with a source.

6 Discussion

The language used in writings about perception often implicitly suggests that perception is typically visual. Thus talk of an “observer” perceiving “objects” or “things” implies vision, possibly also touch. We would not usually talk of observing sounds. However we need to recognize this bias towards vision, and take a more comprehensive view that includes all perceptual modes in our understanding of perception in both natural and artificial domains.

There is often a false dichotomy between competing theories of perception. Taking some of the most prominent features of a sensorimotor account we have looked at whether the perceiving subject is necessarily active. In vision, this is typically the case, but in other perceptual modes such as audition the subject can commonly be either active or passive. This does not mean that the theory of an “active” process in vision is mistaken: the mistake comes in claiming for all types of perception the necessity of a process that does not apply universally.

Another approach is to reconsider what we mean by “active”. The train of thought from Merleau-Ponty and then Varela is that perception requires directed attention, or intentionality. In a contrasting neuroscientific approach the concept of “active” is viewed differently. For instance, Zeki comments on functional specialisation in visual processing - such that different areas of the brain process color, shape, movement autonmously. He says that this “has been instrumental in changing our minds about vision as a process, impelling us to consider it as an active process ....The brain, then, is no mere passive chronicler of the external physical reality, but an active participant ...” [36]. However, this activity in the brain is a subconscious process, not directed or intentional (in either sense). We can refer back to Heidegger’s point that perception can become routine, or subconscious, as in turning a handle to open a door. We may only pay attention if there is some interruption to the usual routine. Now, this type of subconscious activity, in which the action can easily be restored to attention, is one end of a spectrum. At the other end we have neural activity, for instance in integrating the color and shape of an observed object, that we cannot control. In between there are a range of activities that can be moderated with training and effort - for instance the control of a robotic limb discussed above. After a stroke affecting a patient's motor abilities, conscious effort to control muscles that would normally function without thought can be part of a rehabilitation process.

Misconceptions about the sensorimotor account of perception can partly be traced to the ambiguity of the term "representation". It can be taken to mean a reconstructed mental image and a key part of the sensorimotor account is that there are no such images of objects perceived: their own external existence is their representation. However, the term "representation" can in contrast be used to refer to distributed neuronal patterns in the brain, and in this sense all perceptual modes are associated with a representation. Varela et al. discuss the different meanings of "representation" [32]. O'Regan and Noë write of cortical maps [25, page 939] which others might refer to as "representations". These distributed neural patterns are not reconstructed representations like a picture.

A core concept of a sensorimotor account of perception is that the perceiving subject is in the world, not separated by a mediating construction. Our understanding of what constitutes this world is much deeper than a collection of objects that can be seen or touched. It includes auditory perceptions and other perceptual modes found in non-human animals that could inspire robotic development [33]. By deepening our understanding of what constitutes the external world and how we interact with it does not diminish the sensorimotor account of perception, but gives it a firmer empirical base.

References

1. Arena, P., Patane, L.: Spatial Temporal Patterns for Action-Oriented Perception in Roving Robots. Springer (2009), doi:10.1007/978-3-88464-4
2. Baghai-Ravary, L.: Evidence for the strength of the relationship between automatic speech recognition and phoneme alignment performance. In: ICASSP (International Conference on Acoustics, Speech and Signal Processing) (2010)
3. Beeman, W.O.: Linguistics and Anthropology. In: Kempson, R., Fernando, T., Asher, N. (eds.) The Philosophy of Linguistics. Elsevier (2012)
4. Berkeley, G.: Essay towards a New Theory of Vision (1709)
5. Cangelosi, A., et al.: Integration of action and language knowledge: A roadmap for developmental robotics. IEEE Transactions on Autonomous Mental Development 2(3), 167–194 (2010)
6. Curtin, S., Hufnagle, D.: Speech perception. In: Bavin, E. (ed.) The Cambridge Handbook of Child Language. CUP (2009)
7. Greenberg, S.: Speaking in shorthand: A syllable-centric perspective for understanding pronunciation variation. Speech Communication 29, 159–176 (1999)
8. Hatsopoulos, N.G., Donoghue, J.P.: The science of neural interface systems. Annual Review of Neuroscience 249, 249–266 (2009)
9. Heidegger, M.: Being and Time. Blackwell (1927), translated by McQuarrie, J., Robinson, G. published (1962)
10. Hickok, G., Houde, J., Rong, F.: Sensorimotor integration in speech processing: Computational basis and neural organization. Neuron 69(3), 407–422 (2011)
11. Hume, D.: A Treatise of Human Nature. OUP (1740, 1978), book I, part IV, section II
12. Husserl, E.: Logical Investigations. Routledge and K Paul (1900), translated by Findlay, J. published (1970)

13. Hutto, D.: Radically enactive cognition within our grasp. In: Radman, Z. (ed.) The Hand: An Organ of the Mind. MIT Press (2013)
14. Jacquemot, C., Scott, S.K.: What is the relationship between phonological short term memory and speech processing? Trends in Cognitive Sciences 10(11) (2006)
15. Kuhl, P.: Early language acquisition: Cracking the speech code. Nature Reviews - Neuroscience 5, 831–843 (2004)
16. Locke, J.: An Essay Concerning Human Understanding. Thoemmes, published 2003 (1690), book II, chap IX, section 9
17. Lyon, C., Nehaniv, C.L., Saunders, J.: Interactive Language Learning by Robots: the transition from babbling to word forms. PLoS1 7(6) (2012), doi:10.1371/journal.pone.0038236
18. Magee, B.: The Great Philosophers. OUP (1987), reprinted (2009)
19. Merleau-Ponty, M.: Phenomenology of Perception. Routledge, New York (1945), translated by Colin Smith, reprinted (2005)
20. Merleau-Ponty, M.: Husserl at the Limits of Phenomenology. NW University Press (1959), edited by Lawlor, L., Bergo, B. (2002)
21. Näätänen, R.: The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). Psychophysiology 38, 1–21 (2001)
22. Nehaniv, C.L., Förster, F., Saunders, J., Broz, F., Antonova, E., Kose, H., Lyon, C., Lehmann, H., Sato, Y., Dautenhahn, K.: Interaction and Experience in Enactive Intelligence and Humanoid Robotics. In: IEEE Symposium on Artificial Life (IEEE ALIFE). IEEE Symposium Series on Computational Intelligence, SSCI (2013)
23. Noë, A.: Action in Perception. MIT (2006)
24. Noë, A., Thompson, E.: Vision and Mind. MIT (2002)
25. O'Regan, J.K., Noë, A.: A sensorimotor account of vision and visual consciousness. Behavioral and Brain Sciences 24, 939–1031 (2001)
26. Petrini, K., Pollick, F., Dahl, S., McAleer, P., McKay, D., Rocchesso, D., Waadeland, C.H., Love, S., Avanzini, F., Puce, A.: Action expertise reduces brain activity for audiovisual matching actions: An fMRI study with expert drummers. NeuroImage 56, 1480–1492 (2011)
27. Port, R.: How are words stored in memory? Beyond phones and phonemes. New Ideas in Psychology 25, 143–170 (2007)
28. Saffran, J., Aslin, R., Newport, E.: Statistical learning by 8-month-old infants. Science 274, 1926–1928 (1996)
29. Sapir, E.: Le realite psychologique des phonemes. Journal De Psychologie Normale Et Pathologique, 247–265 (1933)
30. Saunders, J., Nehaniv, C.L., Lyon, C.: The acquisition of word semantics by a humanoid robot via interaction with a human tutor. In: Dautenhahn, K., Saunders, J. (eds.) New Frontiers in Human-Robot Interaction. John Benjamins (2011)
31. Tikanoff, V., Cangelosi, A., Metta, G.: Integration of speech and action in humanoid robots: iCub simulation experiments. IEEE Transactions on Autonomous Mental Development 3(1) (2011)
32. Varela, F., Thompson, E., Rosch, E.: The Embodied Mind. MIT Press (1991)
33. Vernon, D., von Hofsten, C., Fadiga, L.: A Roadmap for Cognitive Development in Humanoid Robots. Springer (2010)
34. Wang, R., Perreau-Guimaraes, M., Carvalhaes, C., Suppes, P.: Using phase to recognize English phonemes and their distinctive features in the brain. In: Proceedings of the National Academy of Science, vol. 109(50) (2012)
35. Wolmetz, M., Poeppel, D., Rapp, B.: What does the right hemisphere know about phoneme categories? Journal of Cognitive Neuroscience (2011)
36. Zeki, S.: Inner Vision: An exploration of Art and the Brain. OUP (1999)