ferences. In some circumstances, accurate reality or global reality is not helpful for the animal, as is well illustrated by the tilting train. On a curve when the train cabin is quasi aligned with GI force, the subject has the feeling of being upright while the landscape appears tilted (Neimer et al. 2001). This outside visual flow is a useless referent and severe motion sickness can appear even in a seated reading subject. Closing the blinds to reduce the available information suppresses motion sickness by annihilating the conflicting information provided by the two visual referents (cabin and landscape). Actually, a strong correlation appears between motion sickness triggered by a tilting train and motion sickness induced experimentally during a previous exposition to an oblique rotating optokinetic cone (a control cone rotating in pure yaw exhibits no correlation with the train motion sickness). This implies that some subjects who usually rely on visual geometrics (cabin) and kinematics (outside flow) feel an intrasensorial conflict between referents.

In conclusion, it seems that Stoffregen & Bardy's (2001) heuristic approach to perception may appear, in some circumstances, as unrealistic because of its excessive generality. We agree with the view that each situation is specified by the global array; however, we claim that different animals perceive different subsets of specification. In any case, these differences are piloted by the characteristics of the senses. If we take the Gestalt example used in the target article of the perception of a triangle, animals perceive a part of the whole – that is, they perceive an incomplete triangle which is not an isolated element but a sufficient substructure. This might explain why such theoretical positions as amodality, functional modality, intermodality, and multimodality are sometimes simultaneously possible.

On the subject of perceptual illusions, and the ambiguity of perceptual information

Arve Vorland Pedersen[a] and Hermundur Sigmundsson[b]

[a]Department of Physiotherapy, Sør-Trøndelag University College, N-7004 Trondheim, Norway; [b]Faculty of Social Sciences and Technology Management, Norwegian University of Science and Technology, N-7491 Trondheim, Norway. **Arve.Pedersen@ahs.hist.no Hermundur.Sigmundsson@svt.ntnu.no http://www.svt.ntnu.no/iss/hermundur.sigmundsson/**

Abstract: Stoffregen & Bardy (S&B) make the point that the statement "I am moving" made by subjects in a "swinging room" cannot be explained as an illusion of motion, and there is thus no perceptual illusion. In this they are correct. There is in fact motion, but of the environment. We argue that the subjects misinterpret this because the information to the visual system is ambiguous and also deceiving.

In their target article, Stoffregen & Bardy (S&B) (2001) discuss the frames of reference for motion and use as an example perceptual illusions. They suggest a reinterpretation of subjective reports of physical motion, which are to be found in studies in which perceptual information is manipulated (e.g., Dichgans & Brandt 1978; Lishman & Lee 1973).

The statement "I am moving," reported by subjects in the above-mentioned experiments, is said by S&B to be ambiguous because there is more than one frame of reference. The "swinging room" (Lishman & Lee 1973) or "rotating drum" (Dichgans & Brandt 1978) is one reference frame (the illuminated environment), and the earth is another. S&B say that subjects stating "I am moving" are correct because they are in motion relative to one of the relevant reference frames, namely, the illuminated environment. This interpretation we consider to be wrong – at least when, like S&B, we are not talking about normal postural sway. What is really ambiguous is not the statement "I am moving" but the perceptual information. The problem that subjects encounter in the described situation is that the visual system cannot uniquely

specify motion relative to the environment as self-motion. Gibson (1968) suggested that one possible way for the visual system to distinguish between self-motion and motion of the environment is to detect part-whole differences: "motion perception caused by locomotion entails change in the whole of the textured ambient array whereas the alteration of perspective caused by an objective motion entails only change in part of the ambient array, the remainder being frozen" (p. 187). In the Lishman and Lee (1973) and Dichgans and Brandt (1978) studies, there are no such part-whole differences because the whole room (or drum) is being moved. A change in position relative to the environment, therefore, looks exactly the same whether the change is caused by self motion or is caused by motion of the environment. This is exactly what the subjects encountered. The change in position relative to the environment was caused by motion of the environment and not by any movement, active or passive, of the subject.

When subjects are faced with such a situation, they most often misinterpret this relative motion as self-motion. We argue (cf. Pedersen 1999; 2000) that this is not because – as Lishman and Lee (1973) and later Lee and Aronson (1974) and Lee and Lishman (1975) suggested – the visual system is the dominant system. Neither is it because, as S&B argue, the subjects are in motion relative to the (illuminated) environment, or relative to anything for that matter. It is because the information provided to the visual system is ambiguous, and in this situation, also deceiving.

S&B would, however, seem to be right when they state that this is not a perceptual illusion because there is, actually, motion; therefore, the perception is correct. The subject's visual system does detect motion of the room. The mechanical proprioceptors are also correct because they detect that the body is stationary, and they are not concerned with motion of the room at all. Why would the subject, then, interpret this as self-motion? The problem lies, as mentioned, in the information, which is deceiving because the fact that the room moves is inconsistent with all the subject's experience so far, whether as infant (Lee & Aronson 1974) or as adult (Dichgans & Brandt 1978; Lee & Lishman 1975; Lishman & Lee 1973). This means, in Gibson's (1966) terms, that a genuine biological invariant ("rooms do not move") has been destroyed. In such a situation, says Gibson, a subject must either accept the visual information and reject the postural (mechanical), or accept the postural (mechanical) and reject the visual, or alternate between the two, or compromise between the two. Of course, still according to Gibson, the subject may sometimes just be confused (Gibson 1966, p. 297). In the studies referred to (Dichgans & Brandt 1978; Lishman & Lee 1973), the subjects perceived that something was, in fact, moving, and they reasoned that it could not (at least it should not) be the room (or drum). The information was, therefore, interpreted as propriospecific when it was actually exterospecific.

On ventriloquism, audiovisual neurons, neonates, and the senses

Monique Radeau and Cécile Colin

National Fund for Scientific Research (FNRS) and Université libre de Bruxelles, Brussels, Belgium. **moradeau@ulb.ac.be   ccolin@ulb.ac.be**

Abstract: The analogy between the rules that subtend ventriloquism and bimodal neurons responding suggests a possible neural mechanism for audiovisual interactions in spatial scene analysis. Perinatal data, such as those on synesthesia, sensory deprivation, and sensory surstimulation, as well as neuroanatomical evidence for transitory intersensory connections in the brain support the view that audition and vision are bound together at birth.

As illustration of sensitivity to patterns that extend across the acoustic and optic arrays, Stoffregen & Bardy (S&B) (2001) quote the McGurk effect observed in speechreading (McGurk & McDonald 1976). Additional evidence is provided by the ventrilo-

quism phenomenon that occurs in spatial scene analysis. When light flashes and sound trains come from moderately separated locations, the spatial separation is generally unnoticed, and perceptual recalibration is also manifested in aftereffects (Radeau & Bertelson 1974). The criteria for pairing visual and auditory signals from different locations are sensory factors like the timing of the signals (although strict synchrony is not required) and the distance between them. As demonstrated in barn owls raised from birth to adulthood with prisms (Knudsen & Knudsen 1989) and in human adults (Colin et al. 2001), ventriloquism decreases with increasing spatial separation, being maximal until 20°. Cognitive factors do not play any role. A context simulating a real-life situation, such as seeing the face of a speaker or the hands of a man playing bongos while hearing the sounds displaced, does not enhance ventriloquism beyond the level reached in more artificial situations, as when diffuse light is modulated by the sounds (Radeau & Bertelson 1977) or when the speaker's face is presented inverted (Colin et al. 2001). The system underlying ventriloquism has been considered as being based on primal knowledge of the Gestalt principles of common fate and proximity (Radeau 1994a), used both in visual grouping and in "auditory scene analysis" (Bregman 1990).

Contrary to the ventriloquism effect, which concerns localization, the McGurk effect concerns speech identification and is subtended by different spatial and cognitive rules. It is unaffected by the degree of spatial separation between the signals (Colin et al. 2001), but it decreases in cases of face-voice gender discrepancy for familiar speakers (Walker et al. 1995) or of face inversion (Bertelson et al. 1994; Colin et al. 2001; Jordan & Bevan 1997; Massaro & Chen 1996).

The two effects are probably achieved by specific mechanisms in a way consistent with their different functions (Radeau 1994b). Neurophysiological studies of vision in nonhuman primates have provided evidence for the "what" and "where" problems involving distinct neural pathways (Ungerleider & Mishkin 1982). Recent neuropsychological data from human patients with left hemisphere lesions argued for a "what" versus "where" distinction in the auditory modality as well (Poremba et al. 2003).

The discovery of multimodal neurons helps in understanding crossmodal responses because sensory convergence on individual neurons may well be the underlying neural mechanism. Multisensory neurons have been found in many species and in many parts of the brain. Especially relevant here are the audiovisual neurons found in the deep layers of the superior colliculus and in the polysensory cortex of cat and monkey (Stein & Meredith 1993). Although these neurons often fail to respond to unimodal stimulation, they exhibit vigorous responses under bimodal stimulation provided the stimulations come from locations not too far apart. Enhancement is eliminated around 20° of spatial disparity, and it is inversely related to temporal disparity without being restricted to temporal coincidence. The rules that govern responses of audiovisual neurons are therefore very similar to those that underlie ventriloquism, so these neurons could well constitute the neural substrate of this phenomenon.

What about the development of multisensory functioning? Does it result from amodal representations that are functional early in life or is it learned from experience of co-occurrent unimodal informations, as assumed by empiricist philosophy? Probably due to the immaturity of the superior colliculus of the newborn cat, there is no evidence for multisensory enhancement before several weeks after birth (Stein et al. 2000). However, there is much behavioral evidence to indicate that there is a primitive unity of the senses, the sensory systems becoming gradually differentiated during development (Bower 1974; Gibson 1966; Marks 1978). In the first months after birth, neonatal humans (Lewkowicz & Turkewitz 1980) and rats (Spear & McKinzie 1994) respond to stimulation in all modalities; further, these responses are dominated by quantitative aspects of the stimulation without distinction of modality. On the other hand, synesthesia (joined sensation) is very important in the first month of life and decreases

during development, being two to three times more frequent in infants than in adults (Marks 1975; Maurer 1993).

Data from studies on perinatal sensory surstimulation or sensory deprivation also provide support for early auditory-visual connections. Unusually early experience in a late-developing system interferes with sensory functioning in earlier-developing systems. Exposure of bird embryos to visual stimulation several days prior to hatching results in an auditory deficit, with ducklings (Gottlieb et al. 1989) and quail chicks (Lickliter & Banker 1994) failing to learn the maternal call.

Moreover, perinatal deprivation in a sensory system can affect functioning in the remaining modalities. Deprivation of patterned visual stimulation by binocular eyelid suture in ferrets (King & Carlile 1993) and barn owls (Knudsen et al. 1991) results in anomalous responses of auditory neurons.

Visual event-related potentials (ERPs) have been recorded in congenitally deaf cats (Rebillard et al. 1980) and humans (Neville 1990) over temporal brain areas, which in the hearing subject contain the auditory cortex. However, there was no change in humans who became deaf after the age of four years. Moreover, in congenitally blind humans, auditory and somatosensory ERPs have been found to have a more posterior distribution than in control subjects (Kujala et al. 1992; 1995). The observed compensatory changes can thus reflect stabilization of transitory connections in one modality (Changeux & Dehaene 1989; Edelman 1987) in the absence of competing input from another modality.

There is some neuroanatomical evidence for transient auditory to visual cortex connections around birth that disappear in the fourth week of age in the kitten (Innocenti & Clarke 1984) and in the ferret (Kennedy & Dehay 1993). Connections have also been found between the retinas and the somatosensory and auditory nuclei of the thalamus in the hamster less than 1-week old (Frost 1990). In primate newborns, auditory ERPs have been recorded over the occipital visual cortex of human 6-month-old babies but not in older children (Neville 1995).

All of these data argue for initial sensitivity to structures in the global array, experience probably leading to sensitivity to structures in single-energy arrays.

Retinae don't see

John T. Sanders
Department of Philosophy, Rochester Institute of Technology, Rochester, New York 14623. **jtsgsh@rit.edu**

Abstract: Sensation should indeed be understood globally: some infant behaviors do not make sense on the model of separate senses; neonates of all species lack time to learn about the world by triangulating among different senses. Considerations of natural selection favor a global understanding; and the global interpretation is not as opposed to traditional work on sensation as might seem.

As Stoffregen & Bardy (S&B) (2001) contend, the theory of "direct" perception does indeed seem to be incompatible with the idea that sensory data should be understood as being gathered independently via several senses. In suggesting that the conflict be resolved by rejecting the latter idea, rather than simply by abandoning the theory of direct perception, they choose the more controversial route. In this brief response I offer a few further reasons to take their suggestion seriously.

In the first place, some infant behaviors simply do not make sense on the assumption that at birth the several senses begin providing independent information which can be brought together only after further experience and comparison. Excellent examples