

# Can visual cognitive neuroscience learn anything from the philosophy of language? Ambiguity and the topology of neural network models of multistable perception

Philipp Koralus

Received: 7 April 2014 / Accepted: 21 June 2014 / Published online: 19 July 2014  
© Springer Science+Business Media Dordrecht 2014

**Abstract** The Necker cube and the productive class of related stimuli involving multiple depth interpretations driven by corner-like line junctions are often taken to be ambiguous. This idea is normally taken to be as little in need of defense as the claim that the Necker cube gives rise to multiple distinct percepts. In the philosophy of language, it is taken to be a substantive question whether a stimulus that affords multiple interpretations is a case of ambiguity. If we take into account what have been identified as hallmark features of ambiguity and look at the empirical record, it appears that the Necker cube and related stimuli are not ambiguous. I argue that this raises problems for extant models of multistable perception in cognitive neuroscience insofar as they are purported to apply to these stimuli. Helpfully, similar considerations also yield reasons to suggest that the relevant models are well motivated for other instances of multistable perception. However, a different breed of model seems to be required for the Necker cube and related stimuli. I end with a sketch how one may go about designing such a model relying on oscillatory patterns in neural firing. I suggest that distinctions normally confined to the philosophy of language are important for the study of perception, a perspective with a growing number of adherents.

**Keywords** Ambiguous figures · Nonspecificity · Neural models · Multistable perception · Necker cube · Communication-through-coherence (CTC) hypothesis · Attention

---

P. Koralus (✉)  
Faculty of Philosophy, University of Oxford, Radcliffe Humanities, Woodstock Rd,  
Oxford OX2 6GG, UK  
e-mail: philipp.koralus@philosophy.ox.ac.uk

## 1 Introduction

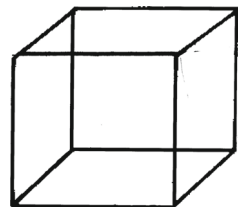
Various static stimuli can give rise to multiple distinct percepts. Well-studied examples are figures like the duck rabbit, the Necker cube (Fig. 1), and similar figures allowing for alternative depth interpretations. Another example is binocular rivalry.

Stimuli engendering perceptual multistability have held enduring interest in the study of the mind and brain for at least 180 years and have been said to offer “a unique window” to the workings of the visual system (Long and Toppino 2004). More recently, both philosophers and cognitive neuroscientists have described stimuli of this kind as important tools for the study of the neural correlates of consciousness (Crick and Koch 1998, 2003; Chalmers 2004). In one particularly interesting sense of the term, “neural correlates of consciousness” would be aspects of brain function that correlate specifically with changes in phenomenal experience and cannot simply be accounted for in terms of changes in sensory stimulation. If we cannot hold sensory stimulation fixed, it is hard to tell whether a particular pattern of neural activity is to be associated with consciousness or merely with sensory processing. Stimuli engendering perceptual multistability potentially allow us to sidestep the confound of sensory stimulation and allow us to study changes in conscious perceptual states independent of changes in perceptual stimulation.

The role perceptual multistability plays in the scientific study of consciousness makes an understanding of the mechanisms underlying perceptual multistability more central than it might appear on a first look. The aim of this paper will be to argue that there are important systematic weaknesses in relevant models that have been proposed in cognitive neuroscience, and to suggest a possible direction to remedy these weaknesses. In particular, I will argue that study of multistable perception would benefit from importing a few concepts from the philosophy of language, in particular, the distinction between ambiguity and non-specificity. The idea that the study of perception may benefit from a semantic perspective, pioneered by Atlas (1989, 2005), to whose insights the present work is heavily indebted, has a growing number of adherents. Work in a similar spirit can be found in Koralus (2010, 2013, 2014), Cumming (1989) and Greenberg (2011), among others. How the findings in this paper affect what conclusions we might want to draw from perceptual multistability about the neural correlates of consciousness will have to be left for another day.

A large number of models of the processes underlying multistable perception have been proposed (See reviews in Borisyuk et al. 2009; Long and Toppino 2004). It has been noted that the empirical record sometimes appears conflicting; some data suggest that low-level mechanisms are responsible for perceptual flips, while other

**Fig. 1** The Necker cube



data are more suggestive of top-down influence (Long and Toppino 2004). I will argue that extant models share an unexamined and often unwarranted assumption about the nature of the representations that underlie different percepts. In short, extant models take as their starting point the assumption that different percepts result from ambiguity. As I will make clear, this assumption is a more specific commitment than the uncontroversial observation that certain stimuli give rise to a variety of different percepts. I will argue that a large productive class of multistable stimuli similar to the Necker cube are not ambiguous and thus are not explained by extant models.

A broader methodological lesson I propose is that vision science may benefit from paying attention to some distinctions familiar in semantic analysis, as done in philosophy, linguistics and, to some extent, computer science. This does not mean that vision science should be conducted from the armchair or rely on “purely conceptual” arguments; the distinctions at issue have concrete empirical consequences, as argued below.

## 2 Models of multistability and the topology of ambiguity

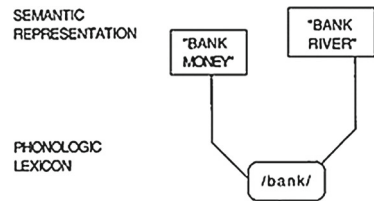
### 2.1 What is ambiguity?

Verbal perceptual inputs like “bank” that have two separate entries in the mental lexicon (e.g. “money bank” and “river bank”), only one of which can be integrated into a sentence at a time, are paradigmatic instances of ambiguity. It is important not to confuse “ambiguous” with “allowing for multiple interpretations”. It has been recognized at least since the thirteenth century that not all stimuli (sentences) that superficially appear ambiguous are in fact ambiguous, and linguists have developed tests for ambiguity (Zwicky and Sadock 1975; Atlas 1989; Ashworth 1991). Some varieties of interpretation are better understood as due to pragmatics (Grice 1975; Atlas 1989). Some of these cases are obvious. For example, nobody would take it that “this philosophy job candidate has excellent handwriting” is ambiguous between “this philosophy job candidate just has excellent handwriting” and “this philosophy job candidate is not qualified for the job”. Other cases are more easily confused with ambiguity. For example, take “the producer started the script”. Is this ambiguous between “the producer started *reading* the script”, “the producer started *writing* the script” and “the producer started *producing* the script”? It is not immediately obvious, but sentences affording multiple interpretations of this sort are better understood as non-specific or “silent” with respect to different interpretations, involving supplementation from our background knowledge, rather than *picking* different mental lexicon entries that store complete different disambiguations wholesale, and there is some evidence that their processing differentially recruits certain brain regions (Pylkkänen and McElree 2007).

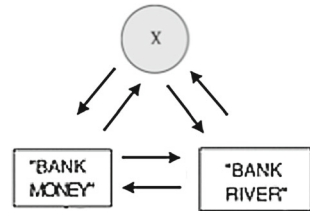
In terms of a network model, ambiguity is a topological feature. As in Fig. 2, separate nodes or sub-networks correspond to separate disambiguations.

At the heart of linguistic tests for ambiguity is the observation that only one disambiguation of an ambiguous constituent like “bank” can be active at a time. As a result, for example, ‘John went to the bank and so did Bill’ is two-ways but not four-ways ambiguous; any literal interpretation of this sentence has to have both John and Bill

**Fig. 2** Adapted from Frost et al. (1990)



**Fig. 3** General pattern of ambiguity networks for “Bank”

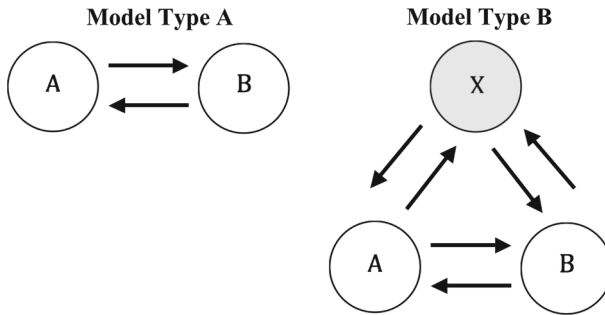


going to a money bank or to a river bank, but not to a mix of the two (Zwicky and Sadock 1975). Ambiguous constituents do not support crossed interpretations (Atlas 1989) (I’ll leave it as an exercise to the reader to show that “started a script” is not ambiguous with respect to the activity started). Thus, the two semantic representations in Fig. 2 have to be in competition for dominance. One possibility would be that the networks corresponding to the semantic representations “Bank Money” and “Bank River” directly inhibit each other. Another possibility is that both networks compete for the resources of some third network “X” that inhibits the loser of the competition. The latter possibility could also coexist in some form with a version of the former, with a combination of inhibitory connections between the “Bank Money” and “Bank River” networks and connections with some third network “X”. All of these possibilities can be represented in an abstract network diagram like Fig. 3, where we can assign different excitatory or inhibitory roles (or similar modulatory roles like gain control) to the different arrows (including the possibility that some arrows have “zero” weights and thus no causal effect).

Thus, Fig. 3 displays in abstract form the topology of neural networks for ambiguous representations.

## 2.2 Ambiguity as a general feature of models of multistability

I will now make the case that neural network models of multistable perception have tended to be models of ambiguity. A wide variety of models have been proposed. As far as I can see, they all have the following features: Each percept type corresponds to a sub-network. These networks compete for dominance. The dominance of one of the sub-networks means that what the sub-network represents corresponds to the current percept. Only one sub-network may be dominant at a time. In addition, there may be an additional “control” network that mediates the competition. Thus, extant neural network models of multistable perception are largely subsumed by the below diagrams, where A and B represent sub-networks each of which underlies one of the



**Fig. 4** Two types of ambiguity networks

two percepts and X represents a possible third “control” network. Individual proposals from the literature can be represented as parameterizations of model types A and B in Fig. 4, as shown in Table 1.

Notice that the general form of a model of perceptual multistability as in Fig. 4 is identical to a model of ambiguity in terms of its network topology: We have two sub-networks representing two interpretations, with the possibility of a third control network, and we have various inhibitory connections between those sub-networks that ensure that only one of the interpretations is dominant. In other words, models of multistable perception that have the form described in Fig. 4 fundamentally embody an ambiguity analysis of multistable perception.<sup>1</sup>

Models of multistable perception vary with respect to the functional role of the arrows, the code for dominance, and the mechanism that leads to shifts in dominance over time. For example, on an early model of type A, the sub-networks A and B both receive excitatory input from the Necker cube stimulus. Units that correspond to one depth interpretation have excitatory connections, while units that correspond to distinct interpretations inhibit each other (Feldman and Ballard 1982). The model can be adapted to different views of what the distributed representations corresponding to each visual interpretation are. With additional assumptions about neural fatigue over prolonged periods of activation, together with random noise in the system, models of this type can predict that we should get uncontrollable switches between the two percepts. The perceptual fluctuations would be explained by successive stages in which one of the two networks becoming fatigued and the other becomes dominant.

To avoid neuroanatomical implausibilities associated with postulating a large number of local inhibitory connections, one proposal has been to adopt a “*k*-Winners Take All” (*k*WTA) model of competition (O’Reilly and Munakata 2000). The key idea is that instead of having individual distributed representations compete by having each unit inhibit its counterpart, one imposes a systemic constraint that only *k* units or fewer may be active at a given time, where *k* is enforced by the influence of some further network X in models of type B. If we add the assumption that units become fatigued over time, we can again explain perceptual alternations over time. Other proposals are

<sup>1</sup> The term “ambiguity” may have other connotations that are not represented in these diagrams, but this is not important for the arguments to follow.

**Table 1** Selection of models as parameters of type A and B

Model	Pattern	Dominance code	AB	BA	XA	XB	AX	BX	Fluct. source
Classical model	Type A	Firing “on”	Inhibitory	Inhibitory					Noise + adapt.
Ogawa et al. (2000)	Type A	Mean firing rate	Inhibitory	Inhibitory					Noise + adapt.
kWTA model	Type B	Firing “on”			Inhibitory	Inhibitory			Noise + adapt.
Moreno-Bote et al. (2007)	Type B	Mean firing rate			Inhibitory	Inhibitory	Excitatory	Excitatory	Noise + adapt.
Klink et al. (2008)	Type B	Mean firing rate	Inhibitory	Inhibitory	Gain control	Gain control			Noise + adapt. + gain control
Borisjuk et al. (2009)	Type B	Synchronization with X			Inhibitory	inhibitory	Excitatory	Excitatory	Plasticity of connections

possible within the same network topology. While there is considerable variation in the assumed type of neural coding and in the details of interactions between units, the majority of mainstream models are essentially variants of type A and type B. Below, I classify what I take to be a representative selection of models as parameters of those diagrams

There are various relative strengths and weaknesses of adopting the various parameters in Table 1, but those are not my primary concern (O'Reilly and Munakata 2000; Leopold and Logothetis 1999; Borsellino et al. 1972). My concern is point out that extant models of multistable perception, though they may vary along many dimensions, are fundamentally ambiguity theories, in the sense in which Fig. 3 displays the general topology of an ambiguity network.

It bears mentioning that there is an interesting gap in the review literature on multistable perception that remains to be filled, and which I cannot attempt to fill here. There is a tendency in the literature to review models of perceptual alternation *in general*, without keeping track of stimuli types used in studies (e.g. Long and Toppino 2004; Leopold and Logothetis 1999). At the time this paper was written, there seemed to be no comprehensive reviews of the past 170 years worth of experimental data on perceptual alternation that systematically classify results by stimuli type. This is not to deny that there are studies that contrast different types of multistable perception (Meng and Tong 2004). Once it becomes clear that it is a substantive hypothesis that a given type of multistable perception is due to an underlying ambiguity, we have to look more carefully at each type of multistability to decide whether that hypothesis is warranted.

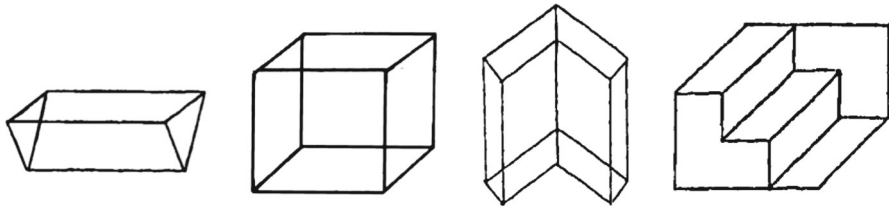
In the rest of this paper, I will argue that for a large productive class of multistable figures like the Necker cube, the ambiguity analysis is mistaken. It should be clear that this is not to say that the models under discussion are useless; I will later consider what types of multistability are plausible cases of ambiguity.

### 3 The representational basis of the Necker cube and related figures

#### 3.1 The representational foundations of models of multistability

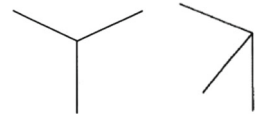
Any neural model of perceptual multistability needs to say what the nodes it postulates represent. In many cases, the focus of the modeling is on the phenomenon of multistability in the abstract, rather than the multistability of a particular class of stimuli. This approach makes it tempting simply to take as a starting point that there are  $n$  percepts, which are going to be represented by  $n$  sub-networks, since we then do not have to bother with details that would reduce the generality of the model. As noted, this move assumes a general hypothesis about the representational underpinnings of multistability, and it is a substantive question whether this yields the right results for all sorts of multistability.

First, I will ask what visual representations are plausibly involved in the percepts we get from figures like the Necker cube. Then, I will examine whether it is plausible that those representations correspond to ambiguity networks of the sort just discussed.



**Fig. 5** Depth-orientation non-specific figures

**Fig. 6** Corner-like line junctions



The instances of multistable perception in Fig. 5 are interesting for only requiring bare contours to generate multiple depth percepts. Contour detection has a special fundamental status in the visual processing hierarchy. Object recognition primarily relies on input from processes that detect contours. Differences between gray-scale photos and line drawings register very little in key brain areas involved in object recognition (Kurtzi and Kanwisher 2000; Hayworth and Biederman 2006). Reaction times for recognition are virtually the same for line drawings and color photographs (Biederman and Ju 1988). Children can identify line drawings by 22–26 weeks of age, and recognition of objects from line drawings is possible even if no pictures of any sort were previously encountered (Hochberg and Brooks 1962; Yonas et al. 1978). In sum, the processing of line-based stimuli bears on a core aspect of visual processing.

How the visual system detects (flat) contour segments of various orientations when presented with the sorts of stimuli illustrated above is about as well understood as any subject in visual cognitive neuroscience. Networks in primary visual cortex can represent oriented segments at various points in retinal space via a population code over various neurons that selectively respond to luminance gradients of different orientations. What is less clear is how a depth impression is created from figures like the Necker cube.

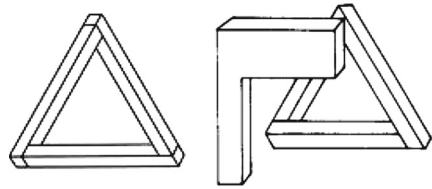
### 3.2 Corners as depth cues

First, we have to ask what visual representations are plausibly involved in the percepts we get from the Necker cube and related figures. Then, we must examine whether it is plausible that those representations give rise to ambiguity networks of the sort just discussed. We are dealing with a systematic class of figures, so the representations are unlikely to be holistic. What features do all those line-based figures that give rise to multiple depth interpretations have in common? Line junctions that could *roughly* be the projections of cubic corners, as in Fig. 6.

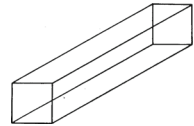
One way to argue that the way we assign depth interpretations to figures based on line-junctions like the above is not fundamentally holistic is to find cases in which we assign depth interpretations yielding apparent object configurations that we could



**Fig. 7** Adapted from [Rock \(1983\)](#)



**Fig. 8** Adapted from [Gillam \(1979\)](#)



never encounter in the real world because they are impossible. We perceive the left in Fig. 7 as presenting a solid object, even though the object visually presented evidently could not be a real object in Euclidian three-dimensional space.

Impossible figures support the idea that the relevant depth cues operate locally. The visual system does not provide a percept of the Penrose Triangle that is geometrically possible, even though the Penrose Triangle could in fact (roughly) be the projection of a slightly unusual real object viewed from a certain angle. In the right in Fig. 7 one of the corners is occluded, and a coherent percept is more readily obtained ([Rock 1983](#)).

The local nature of the mechanisms that suggest depth in line drawings of this sort is evident in even simpler figures.

Most observers spontaneously perceive Fig. 8 as an impossible figure, with both ends of the rectangular box facing the viewer, even though there is a plausible visual interpretation that would be spatially coherent ([Gillam 1979](#)). Again, these figures suggest that depth interpretations of line drawings are computed locally without decisive global coherence constraints. Thus, it is plausible that the corner-like line junctions the figures under considerations have in common trigger a depth cue that is a local feature.

Why would we have a depth cue of this sort based on monocular contour configurations? For vision beyond 30 m of distance from our retinas, we need to rely on monocular depth cues ([Ware 1995](#)). This leaves open what aspect of the scene we should treat as depth cues. One consideration is that two-dimensional projections leave distal object layouts underdetermined. Yet, it is an interesting fact that with minimal assumptions, projections of corners enable good approximations of three-dimensional edges in the distance ([Perkins 1968](#); [Mulder and Dawson 1990](#)). Thus, it would make geometric sense to have detectors for corner-like line junctions that are processed as depth cues.

There is evidence from electroneurophysiology that there are neurons in primary visual cortex as well as in inferotemporal cortex that are selectively sensitive to the sorts of line junctions that could correspond to corners in depth ([Shevelev et al. 2001](#); [Tanaka 1996](#)). In perceptual psychology, it has been argued that rapid search performance that is only linearly dependent on the number of distractors is diagnostic of primitive visual features ([Treisman and Gelade 1980](#)). There is some evidence that line junctions are in fact detected rapidly in this way ([Enns and Rensink 1991](#)). One reason to

think that the visual system treats line junctions of the relevant sort as depth cues is that naïve observers report perceiving depth even in isolated line junctions that could correspond to corners (Perkins 1971, 1972; Shepard 1981, 1990). There is also evidence that line junctions that correspond to corners rather than surface markings exhibit salience for observers as young as 7.5 months (Yonas and Arteberry 1994). Finally, the propensity to perceive depth in simple line drawings with corner-like line junctions is present even in Bushmen of the Northern Kalahari growing up in non-carpentered environments (Deregowski and Bentley 1986; Deregowski 1989). On the basis of observations of this sort, it has been suggested that the basic ability to see corner-like line junctions in depth is largely independent of experience with geometrically precise corners (Deregowski and Bentley 1986; Deregowski 1989). If detection of corners in depth is indeed performed in early visual processing, this is not surprising. The core architecture of low-level visual feature detectors appears to be arranged before any input from visual experience. Even before natural eye opening, orientation selective cells can be found in newborn kittens (Crair et al. 1998). The gross mapping of different orientations onto primary visual cortex seems to stay largely the same through maturation, unless the animals grow up with abnormal visual experience (Crair et al. 1998; Chapman et al. 1996; Blakemore and Cooper 1970).

In sum, it appears plausible that the visual system has local representations of corner-like junctions that serve as depth cues. Evidence from electrophysiology, perceptual psychology, and cross-cultural psychology suggests that line junctions that are perceived as corners are visual coding primitives. There also appear to be good reasons to suspect that those line junctions are interpreted as primitive depth cues. It appears that the visual system includes dedicated corner-in-depth (CD) detectors. Trehub (1991) came to a similar conclusion. This is plausibly what underlies the possibility of depth percepts from figures like the Necker cube. The question is what exactly these CD detectors encode with respect to depth orientation.

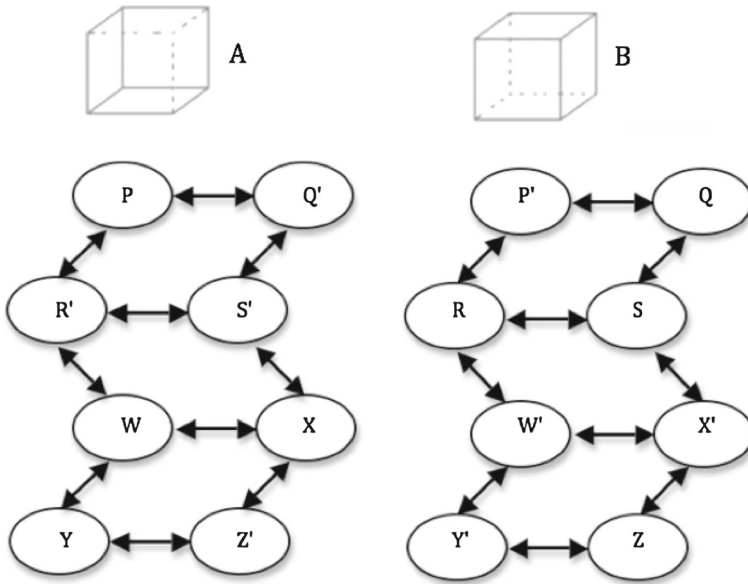
## 4 Against the ambiguity of CD detectors with respect to depth orientation

### 4.1 The ambiguity hypothesis

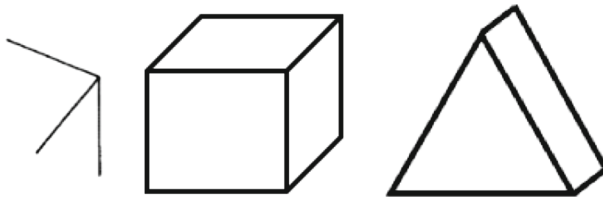
One possibility would be that the corner representations are ambiguous with respect to depth orientation. This brings us back to the network topology of the models of multistable perception considered earlier. We would effectively have two units for every corner detector, one representing a convex corner (represented by a letter in the diagram below) and a corresponding one representing a concave corner (represented by a primed letter). There are several observations that speak against this analysis (Fig. 9).

### 4.2 Redundancy and stability

It is a familiar thought in the study of semantic representations in language that one should not postulate ambiguous representational constituents unless absolutely necessary, often referred to as Grice's Modified Occam's Razor (Grice 1975). If cubic



**Fig. 9** The ambiguity hypothesis

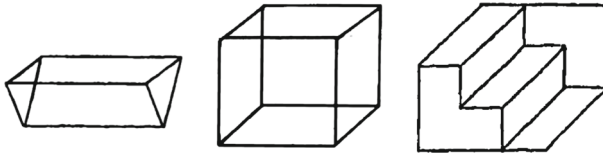


**Fig. 10** Stable depth-orientation non-specific figures

corner representations are ambiguous, we need at least double the number of nodes devoted to corners. If those representations are local ones in primary visual cortex, then this means a lot of extra neurons.

A further worry is that if the multiple depth interpretations of the Necker cube are due to competition between convex and concave corner detectors, any line layouts seen in depth due to these corner detectors would be expected to produce qualitatively the same phenomenon of multistability. However, it is not at all clear that simpler figures like the below give rise to the same sort of multistability. In contrast to those in Fig. 11, it is quite hard to perceive the corners in Fig. 10 as anything other than convex. At the same time, there are no perspectival cues that rule out that would rule out an alternative depth interpretation.

The multistability in Fig. 11 contrasts with the relative stability in Fig. 10. On the ambiguity hypothesis about CD detectors, it is not clear why this should be so.



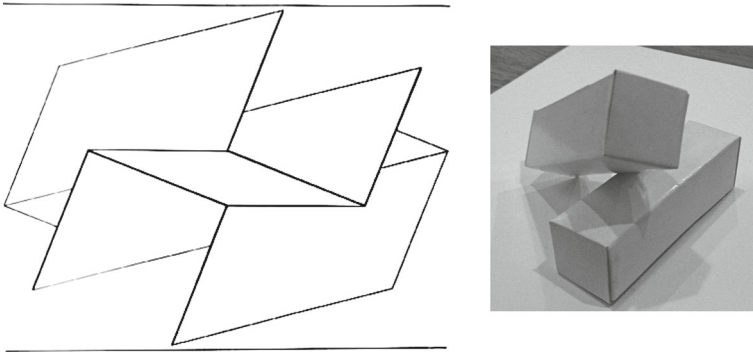
**Fig. 11** Unstable depth-orientation non-specific figures

### 4.3 Crossed depth percepts

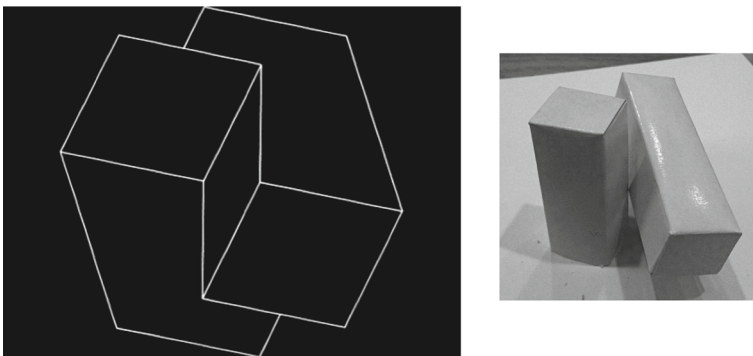
Is it possible to find more direct perceptual evidence for the claim that CD detectors are not ambiguous with respect to depth orientation? In linguistics and philosophy, it has been observed that a paradigmatic feature of ambiguous representations is that only one disambiguation can be integrated into an interpretation or “verbal percept” at a time. Thus, “John went to a bank” is two-ways ambiguous, since “bank” is two-ways ambiguous, but “John went to a bank and so did Bill” is not four-ways ambiguous. “Bank” cannot do double duty to be simultaneously interpreted in two different ways with respect to its possible disambiguations (Zwicky and Sadock 1975). These sorts of constraints do not hold for differences in interpretation that are only due to supplementation from context (Atlas 2005, 1989). A similar idea underlies the ambiguity model of multistable perception: two networks are in a winner-takes-all competition and the network that loses does not contribute its representational content at all. Thus, if the contribution of CD detectors is ambiguous with respect to depth orientation, then insofar as you perceive them in depth at all (without shading or binocular disparity cues) you perceive them as convex or concave but not both at the same time. Now, one might think that of course you cannot perceive the same corner as both convex and concave because no real corner could possibly be both. However, as the impossible figures discussed above made clear, possibility in the distal world is not a decisive constraint on perceptual possibility.

Rather surprisingly, it is in fact possible to perceive line layouts corresponding to corners as doing double duty with respect to convex and concave interpretations. Some of the best demonstrations are due to the Bauhaus artist (and former Dean of Yale’s School of Architecture) Joseph Albers. Discussing works in his *Structural Constellation* series, he remarks on the possibility of perceiving the center parts of the figures as “simultaneously in a forward and a backward direction” (Albers 1977). Consider Figs. 12 and 13 with respect to the spatial configurations indicated in the photos on the right of each drawing.

One can perceive Figs. 12 and 13 as presenting spatial configurations roughly like those indicated in the accompanying photos. On each of those percepts, at least two of the three-line junctions in the center of the figure correspond to both convex and concave corners on the object configuration that is perceived. Those three-line junctions do *double duty* as convex and concave, giving the figures an illusory quality. Note that “able to perceive” does not mean “will always perceive”. For example, if we strongly focus on the center of the figures, the unusual visual interpretations tend to disappear. Attentional focus has a tendency to make things appear as in the foreground, if this is a perceptual possibility afforded by a line display, so it has to be kept in mind



**Fig. 12** Albers (1977). Structural constellation VI. Princeton University Art Museum



**Fig. 13** Albers (1977). Structural constellation

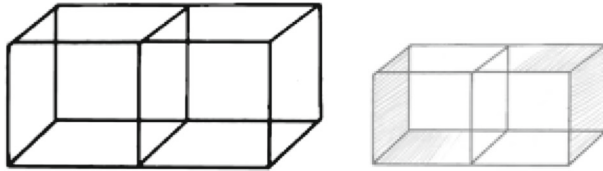
that in cases of displays like the above, “scanning” the image with attentional focus in fact alters the percept (Kawabata 1987; Kawabata and Yamagami 1978; Peterson and Gibson 1991).

The first suggestion that the Necker cube is not ambiguous because it allows for crossed depth interpretations is due to Atlas, who discovered that two Necker cubes drawn in such a way that they share corners, can still be perceived in different depth orientations (even if the normal tendency is to perceive both in the same depth orientation) (Atlas 1989, 2005). See Fig. 14.

The fact that it is possible to perceptually integrate a line junction as a convex corner and as a concave corner simultaneously strongly suggests that the ambiguity theory is the wrong account of the representational contribution of corner detectors.

#### 4.4 Learning

If the ability to see depth in line drawings like the Necker cube were based on convex/concave ambiguous representations then being able to perceive depth in any relevant drawing should predict the ability to get *some* depth interpretation for all



**Fig. 14** Adapted from Atlas (2005)

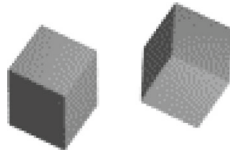
others, even though those interpretations may sometimes be implausible or incoherent. On an ambiguity account of the contribution of CD detectors with respect to depth, each corner representation by itself suggests either a convex or a concave corner.

However, it appears that though all children appear to be able to obtain depth percepts from some corner-based line drawings, the ability to obtain them for more complex ones is subject to learning, even if we allow impossible depth configurations (Young and Deręgowski 1981; Deręgowski 1969; Deręgowski and Bentley 1986). Deręgowski and Dziurawiec (1986), reviewing developmental and cross cultural evidence, suggests that the basic ability to use corner-like line junctions as depth cues is either innate or acquired without need for exposure to carpentered environments, while the ability to obtain a depth interpretation in more complex line layouts using the same cues involves learning (Deręgowski 1989). On the ambiguity hypothesis, this is unexpected.

#### 4.5 Corner depth-cues as orientation non-specific

The foregoing observations suggest that corner representations are not ambiguous with respect to depth orientation. As noted, a flexible visual system needs monocular depth-cues, and geometrical facts make line junctions that likely correspond to corners a good candidate for a feature detector that would serve as a depth cue. However, the advantages of using such line junctions as depth cues do not extend to settle the question whether the corner in the distance is convex or concave. If we just consider geometrical constraints, it would be a good solution to let a visual system include detectors for line junctions that likely correspond to corners. If this detector is only interpreted as signaling what geometry helps it estimate, this detector should be taken as leaving open whether the corner is convex or concave. From an informational perspective, the best solution may be to let those corner representations represent depth, but not depth orientation with respect to the viewer.

In what way could orientation-nonspecific corner representations be useful to the visual system? Major theories of object recognition are based on the neurophysiologically well-grounded view that higher-level visual processing areas like IT detect features of objects of an intermediate level of complexity. On an influential view, the way in which objects are represented is largely invariant to orientation with respect to the viewer (Biederman 2001). An important observation is that even pigeons seem to spontaneously categorize “box like” stimuli in a way that is relatively independent of perspective, regardless of whether they are presented as viewed from above or below.



**Fig. 15** Adapted from Peissig et al. 2002

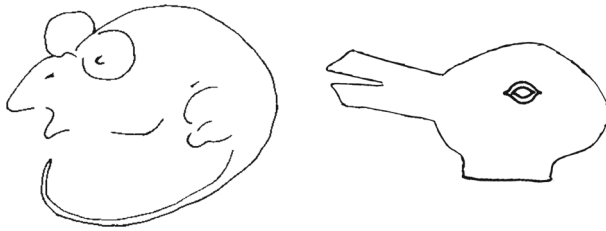
Peissig et al. trained pigeons to peck four different buttons in response to four different shapes. Pigeons that were trained to peck a certain button when presented with the left figure in Fig. 15 reliably pecked that button in response to the right figure (Peissig et al. 2002).

If recognition of objects of this sort is largely viewpoint independent, it would be unnecessary and perhaps unhelpful for CD detectors to encode particular orientations. For ventral-stream visual processing aimed at classifying objects, orientation non-specific corner detectors may be a good solution.

#### 4.6 Contrasting more plausible cases of ambiguity

There are important differences between the duck-rabbit and similar “imagistic” ambiguous figures on the one hand and the Necker cube and related figures on the other. For example, there is greater possibility for control of perceptual fluctuations for the duck-rabbit compared to the Necker cube. It is both easier to slow fluctuations and to speed them up (Strüber and Stadler 1999). There is also a greater possibility for controlling perceptual fluctuations with the Necker cube than with binocular rivalry (Meng and Tong 2004). This provides some support for the notion that the mechanisms underlying perceptual fluctuation in the case of the Necker cube are different. I argued that it is implausible that an ambiguity network underlies perceptual fluctuation in the case of the Necker cube. However, it seems quite plausible to postulate such a network for figures like the duck-rabbit and for binocular rivalry. A plausible theory of object recognition would most likely include separate units for ducks and rabbits, and the demands of object classification might even independently motivate inhibitory connections between those units. In the case of the leftmost in Fig. 16 that can be perceived as a mouse or a face, it is even more plausible that different percepts correspond to separate neural networks, since it has been argued that there is a dedicated area of IT, the fusiform face area (FFA), that selectively processes faces but not animal pictures (Kanwisher et al. 1999; McCone et al. 2007). As for binocular rivalry, we know that cortical columns in primary visual cortex differ in whether they are primarily activated by the left or right eye (Blasdel et al. 1995). Thus, there is independent reason to postulate distinct neural networks corresponding to left-eye and right-eye dominant percepts.

In sum, though the Necker cube and related figures do not appear to be ambiguous, there are plenty of other types of multistable perception that are plausibly seen as cases of perceptual ambiguity.



**Fig. 16** Ambiguous figures

## 5 Toward a neural network without ambiguity: the NAPS (non-accidental property constrained synchronization) model

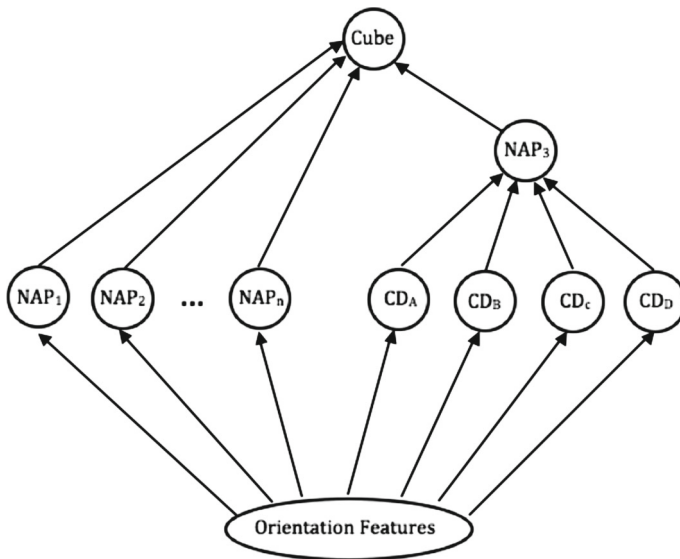
I will now sketch how one might construct a new model of the Necker cube that does not have the topology of an ambiguity analysis. The main aim for this section is to make the case that the negative arguments in the previous sections can in fact serve as a new starting point for a constructive project of designing non-ambiguous models of perceptual multistability.

### 5.1 NAPs and object recognition

My starting point is a rough model of object recognition for figures like the Necker cube. Some theories of object recognition rely on viewpoint-dependent “templates”, others rely on collections of metric features. Still others rely on detecting combinations of non-accidental properties (NAPs) like symmetries, curvature and line intersections that do not change much under changes in viewpoint (Biederman 2001). NAPs are features of an image that are unlikely to be a consequence of an accident of viewpoint and that are highly likely to have corresponding properties in the object itself (Lowe 1984; Witkin and Tenenbaum 1983). For example, if line segments in a retinal projection are collinear, the corresponding edges in three dimensions are likely collinear as well. Similarly, the symmetries and parallels in a projection of an object are likely mirrored in the object itself. Importantly, NAPs are detected as properties of the 2D projection; they do not rely on depth interpretation (Biederman 1987). Data from visual search tasks as well as electrophysiological evidence suggests that there are indeed detectors for NAPs in the visual system (Wolfe and Friedman-Hill 1992; Vogels et al. 2001; Kayaert et al. 2003, 2005). It is likely that both sorts of accounts are required for the full range of objects we can recognize. As noted, recognition for cube-like objects seems to be largely invariant to orientation, so a model that relies on NAPs for recognition is plausible as core model for the type of figure under consideration (Peissig et al. 2002).

It is plausible that among the NAP features that are involved in detecting a cube are certain arrangements of corners. It is also plausible that relevant processing is done in a feed-forward manner, as has been argued is the case for rapid object identification (Serre et al. 2007). This would yield a network layout like the below. Proposals for the neural transfer functions for a network as in Fig. 17 are readily available (*ibid*).





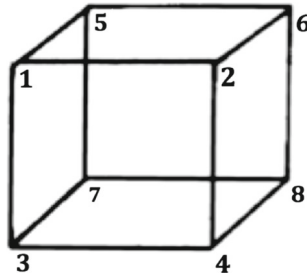
**Fig. 17** NAPs network

The above sketch of a neural network model does not fall into the ambiguity pattern. There are no nodes that are unique to one depth interpretation or another, because *so far* nothing encodes depth. We now have to address how this network could encode additional information about depth orientation without interfering with the feed-forward processing involved in object recognition and without falling back into the mold of the ambiguity models criticized above. I propose that differences in depth orientation are encoded by different patterns of oscillatory coherence.

## 5.2 Depth organization via patterns in neural oscillation

For the purposes of the Necker cube in the simplified framework proposed, we only need to encode which parts of the figure are in the foreground and which are in the background. We could say that if two features are both in the foreground, they both fall into a certain pattern of neural oscillation. Now, we want the detection of a cube to serve as a constraint on which depth interpretations remain possible for parts of the Necker cube. If, say, the corner detectors corresponding to line junctions 1 and 2 do not have similar patterns of oscillation, then signals originating from these detectors will not arrive at the same time at the relevant NAP detectors that would allow us to recognize a cube. The result may be that the relevant NAP is not detected and thus we do not see a cube. Insofar as we detect a cube, what corresponds to nodes 1 and 2 in Fig. 18 may have to be in synchrony and thus has to be given the same depth interpretation.

On this view, *given* a certain set of active NAP feature detectors, certain combinations of depth interpretations are ruled out.



**Fig. 18** Numbered nodes

The underlying mechanism that would account for this constraint relies on the communication-through-coherence (CTC) hypothesis, according to which the effectiveness of communication between neuronal units is proportional to the degree of oscillatory coherence between the units (Fries 2000). At certain points within the oscillation cycle of the neuron, units are more excitable. If a spike from one unit arrives within the excitable window of the receiving unit, communication is possible. If such a spike arrives within the window of least excitability, communication is inhibited.

If a depth interpretation puts units feeding a crucial NAP into sufficiently different oscillatory patterns, it may inhibit detection of that NAP. If we continuously perceive a cube, we may not be able to assign a depth interpretation that breaks up NAPs for this reason. Thus, an object recognition network can constrain depth interpretation, even if depth is not encoded through “depth nodes” in the recognition network.

In principle, it is possible to have multiple added oscillatory patterns at the same neural units, and so, in principle, the proposed framework does not rule out crossed depth percepts, as long as the relevant oscillatory patterns allow for all NAPs involved in seeing the crossed-depth configuration to get adequate input. In sum, the core idea of the proposed model is NAP-constrained synchronization of neural activity (NAPS). This NAPS model can in principle allow for the sorts of crossed-depth percepts that seemed to cast doubt on ambiguity models.

On the NAPS model, the alternating depth interpretations of the Necker cube can be explained using exactly the same network of units we need to understand cube recognition. The only modification is that the connections not only have to support feed-forward integrate-and-fire processing, but also synchronized oscillatory activity. It is usually thought that neural synchrony in the brain relies on reentrant connections, so we would at least add “downward” arrows to the diagram in Fig. 16. Since we want object recognition to constrain depth interpretation, we would predict two processing stages. A first pass, feed-forward stage without a special synchronization pattern that culminates in cube detection is the first stage. The second stage establishes a pattern of neuronal oscillations that corresponds to a particular depth interpretation, with the constraint that the activation that underlies recognition has to remain essentially undisturbed.

What would provide the initial impetus for the visual system to begin assigning an oscillatory pattern to the network that would yield one depth interpretation rather than

another? I suggest that attentional focus provides the seed for establishing this pattern. The model I will sketch has some similarities to the model proposed by [Trehub \(1991\)](#).

### 5.3 The role of attentional focus

There is considerable evidence to suggest that shifts in attention, some voluntary and some involuntary, drive changes in perceived depth orientation. In an unbiased Necker cube, the corner nearest ocular fixation is very likely to be seen as convex and in the foreground ([Kawabata and Yamagami 1978](#)). Adding bold lines or other biases that attract attention near a corner makes it much more likely that that corner will be seen as in the foreground at first ([Kawabata 1987](#); [Peterson and Gibson 1991](#)). Finally, the perceived depth orientation of a Necker cube spontaneously alternates over time, but if participants are told to hold their attention fixed as much as they can, the alternation rate slows down ([Meng and Tong 2004](#)). Diverting observers' attention has a similar effect ([Reisberg and O'Shaughnessy 1984](#)). Though shifts in attention are normally tracked by shifts in ocular fixation, the latter is not necessary for the former. It is widely accepted that "covert" attention shifts independently of eye focus are possible ([Wright and Ward 2008](#)). Correspondingly, retinally fixed Necker cubes still allow perceptual flips ([Pheiffer et al. 1956](#)).

It is a peculiar feature of visual attention that it has both stimulus-driven aspects that require no conscious involvement, and central goal-driven aspects as well ([Wright and Ward 2008](#)). This makes attention seem like a good candidate for explaining the peculiar mix of uncontrollability and partial control involved in the Necker cube. With instructions to speed up perceptual reversals, subjects can increase the frequency of perceptual flips of the Necker cube, as well as lower it ([Strüber and Stadler 1999](#)). Since it is possible to consciously influence the deployment of visual attentional focus, this is not a surprise on the proposed model.

Yet, over prolonged inspection, the perceived orientation of the Necker cube flips spontaneously. The frequency of perceptual flips can be consciously influenced but is not subject to decisive conscious control ([Strüber and Stadler 1999](#); [Long et al. 1983](#); [Babich and Standing 1981](#)). Dominance durations of Necker cube percepts are sequentially stochastically independent and gamma distributed ([Borsellino et al. 1972](#)). Now, one of the functions of attentional focus is to provide more precise analyses of parts of the visual scene. This means that attention should normally wander after analysis of a given part of the visual scene is complete. Moreover, attentional wandering should be "anarchic" under free-viewing conditions, to ensure faster scan times, instead of waiting for a conscious decision to shift attention ([Wolfe 2000](#)). Supporting the view that how we perceive the Necker cube is mediated by attentional focus, just as dominance durations of Necker cube percepts are stochastically independent and gamma distributed under free-viewing conditions, uncontrolled shifts in attention are stochastically independent and gamma distributed as well ([Harris et al. 1988](#); [Suppes et al. 1983](#); [Richards and Gibson 1997](#); [Leopold and Logothetis 1999](#)). In sum, it seems plausible that the perceptual dynamics of the Necker cube is at least initially driven by the dynamics of attentional focus assignment.

Does this picture of attentional focus being involved in beginning to establish an oscillatory pattern that would yield a depth percept fit with neural data? There is evidence that involuntary attention shifts correlate with shifting patterns of relatively lower-frequency synchronization (Vidal et al. 2006). Interestingly, it has also been reported that synchronous patterns of activity around 8–14 Hz spontaneously fluctuate in visual cortex, and that spatially specific activity co-varies with whether a stimulus will be more readily detected (Romei et al. 2008). This would not be surprising if it is the focus of attention that is reflected in patterns of lower-frequency oscillation, since attentional focus has been known since Helmholtz to provide an advantage for stimulus detection. If those patterns play the role suggested, we should find that increased perceptual flipping frequency should decrease the amount of synchronization in lower-frequency bands, as patterns of synchronization corresponding to a particular depth interpretation are more frequently broken up. It has in fact been reported that during continuous viewing, a higher number of perceptual flips correlates with greater desynchronization in the lower alpha band (6–8 Hz) around the posterior area (Isoglu-Alkac and Strüber 2006). In a different study that compared the effect of speeding up and slowing down fluctuations, it was found that delta band (0–4 Hz) synchronization was maximal for the hold condition and minimal for the speed condition (Mathes et al. 2006).

Furthermore, there is evidence that visual focal attention shifts recruit parietal cortex, particularly right parietal cortex (Vidal et al. 2006; Corbetta et al. 1995). On the proposed model, reversing the Necker cube percept characteristically involves a shift in attentional focus. As expected, reversal frequency is decreased in patients with right hemisphere lesions, who also have trouble with visual search that requires attention shifts (Cohen 1959). Increased positivity in the right inferior parietal cortex seems to precede reports of perceptual reversals of the Necker cube (Britz et al. 2009).

In sum, synchronous activity in the 0–8 Hz range seems to correspond both to the maintenance of particular depth interpretation of the Necker cube, and activity in this range has been independently identified as a correlate of the focus of involuntary attention. This suggests that the 0–8 Hz band in fact roughly exhibits some of the properties we expect from the type of pattern of neural oscillation postulated by the NAPS model.

We do not have to assume that a particular pattern of neural synchrony on a certain bandwidth itself has the representational content of special foreground or background. The same broad region of the brain that has been argued to have a key role in attention also includes representations of egocentric spatial information (Seubert et al. 2008). A certain oscillatory patterns might encode depth through synchronization with nodes in parietal lobe that might more generally represent foreground relative to the viewer.

## 6 Conclusion

In sum, I have argued that the Necker cube is not ambiguous, which casts doubt on a wide range of neural models of multistable perception. I sketched how one may go about developing a novel kind of neural network model of the underlying perceptual processes based on the idea of NAP constrained patterns of neural oscillation. A

broader methodological upshot of the proposed analysis is that questions about the precise representational content of coding primitives matter for cognitive neuroscience just as they matter for linguistics and the philosophy of language. In sum, there is work to be done for semanticists in the study of perception.

## References

- Albers, J. (1977). *Despite straight lines*. Cambridge: MIT Press.
- Ashworth, E. J. (1991). Signification and modes of signifying in thirteenth-century logic: A preface to aquinas on analogy. *Medieval Philosophy and Theology*, 1, 39–67.
- Atlas, J. (1989). *Philosophy without ambiguity*. Oxford: Clarendon Press.
- Atlas, J. (2005). *Logic, meaning, and conversation*. Oxford: Oxford University Press.
- Babich, S., & Standing, L. (1981). Satiation effects with reversible figures. *Perceptual and Motor Skills*, 52, 203–210.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115–147.
- Biederman, I. (2001). Recognizing depth-rotated objects: A review of recent research and theory. *Spatial Vision*, 13, 241–253.
- Biederman, I., & Ju, G. (1988). Surface vs. edge based determinants of visual recognition. *Cognitive Psychology*, 20(1), 38–64.
- Blakemore, C., & Cooper, G. F. (1970). Development of the brain depends on the visual environment. *Nature*, 228, 477–478.
- Blasdel, G., Obermayer, K., & Kiorpes, L. (1995). Organization of ocular dominance and orientation columns in the striate cortex of neonatal macaque monkeys. *Visual Neuroscience*, 12, 589–603.
- Borisyuk, R., Chik, D., & Kazanovich, Y. (2009). Visual perception of ambiguous figures: Synchronization based neural models. *Biological Cybernetics*, 100, 491–504.
- Borsellino, A., De Marco, A., Allazetta, A., Rinsei, S., & Bartolini, B. (1972). Reversal time distribution in the perception of visual ambiguous stimuli. *Kybernetik*, 10, 139.
- Britz, J., Landis, T., & Michels, C. M. (2009). Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cerebral Cortex*, 19, 55–65.
- Chalmers, D. (2004). How can we construct a science of consciousness? In M. Gazzaniga (Ed.), *The cognitive neurosciences III*. Cambridge: MIT Press.
- Chapman, B., Stryker, M. P., & Bonhoeffer, T. (1996). Development of orientation preference maps in ferret primary visual cortex. *Journal of Neuroscience*, 16(20), 6443–6453.
- Cohen, L. (1959). Perception of reversible figures after brain injury. *Archives of Neurology and Psychiatry*, 81, 765–775.
- Corbetta, M., Shulman, G. L., Miezin, F. M., & Petersen, S. E. (1995). Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science*, 270, 802–805.
- Crair, M. C., Gillespie, D. C., & Stryker, M. P. (1998). The role of visual experience in the development of columns in cat visual cortex. *Science*, 279, 566–570.
- Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, 8, 97–107.
- Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, 6(2), 119–126.
- Cumming, S., (ms). (1989). The attentional foundations of coherence. Manuscript, UCLA.
- Deregowski, J. (1969). Perception of the two-pronged trident by two- and three- dimensional perceivers. *Journal of Experimental Psychology*, 82, 9–13.
- Deregowski, J. (1989). Real space and represented space: Cross-cultural perspectives. *Behavioral and Brain Sciences*, 12, 51–119.
- Deregowski, J., & Bentley, A. M. (1986). Perception of pictorial space by Bushmen. *International Journal of Psychology*, 21, 743–752.
- Deregowski, J., & Dziurawiec, S. (1986). Some aspects of comprehension of technical diagrams: An intercultural study. *Le Travail Humain*, 49, 43–60.
- Enns, J. T., & Rensink, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review*, 98(3), 335–351.
- Feldman, J. A., & Ballard, D. H. (1982). Connectionist models and their properties. *Cognitive Science*, 6, 205–254.

- Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10), 474–479.
- Frost, R., Feldman, L. B., & Katz, L. (1990). Phonological ambiguity and lexical ambiguity: Effects on visual and auditory word recognition. *Journal of Experimental Psychology*, 16, 569–580.
- Gillam, B. (1979). Even possible figures can look impossible. *Perception*, 8, 229–232.
- Greenberg, G. (2011). The Semiotic Spectrum. Doctoral dissertation. Rutgers University.
- Grice, H. P. (1975). Logic and conversation. In J. Kimball (Ed.), *Syntax and semantics* (Vol. 3, pp. 139–151). New York: Academic Press.
- Harris, C. M., Hainline, L., Abramov, I., et al. (1988). The distribution of fixation durations in infants and naïve adults. *Vision Research*, 28, 419–432.
- Hayworth, K. J., & Biederman, I. (2006). Neural evidence for intermediate representations in object recognition. *Vision Research*, 46, 4026–4031.
- Hochberg, J. E., & Brooks, V. (1962). Pictorial recognition as an unlearned ability. *American Journal of Psychology*, 75, 624–628.
- Isoglu-Alkac, U., & Strüber, D. (2006). Necker cube reversals during long-term EEG recordings: Sub-bands of alpha activity. *International Journal of Psychophysiology*, 59(2), 179–189.
- Kayaert, G., Biederman, I., Op de Beeck, H., & Vogels, R. (2005). Tuning for shape dimensions in macaque inferior temporal cortex. *European Journal of Neuroscience*, 22, 212–224.
- Kayaert, G., Biederman, I., & Vogels, R. (2003). Shape tuning in macaque inferior temporal cortex. *Journal of Neuroscience*, 23, 3016–3027.
- Kanwisher, N., Stanley, D., & Harris, A. (1999). The fusiform face area is selective for faces not animals. *NeuroReport*, 10, 183–187.
- Kawabata, N. (1987). Interpretive process of depth in line drawing. *Systems and computers in Japan*, 18(7), 103–109.
- Kawabata, N., & Yamagami, K. (1978). Visual fixation points and depth perception. *Vision Research*, 18, 853–854.
- Klink, P. C., van Ee, R., Nijs, M. M., Bruwer, G. J., Noest, A. J., & van Wezel, R. J. A. (2008). Early interactions between neuronal adaptation and voluntary control of perceptual choices in bistable vision. *Journal of Vision*, 8(5), 1–18.
- Koralus, P. (2010). Semantics in Philosophy and Cognitive Neuroscience. PhD Dissertation. Princeton University.
- Koralus, P. (2013). Attention, Consciousness, and the semantics of questions. *Synthese*, doi:10.1007/s11229-013-0382-1.
- Koralus, P. (2014). The erotetic theory of attention: Questions, focus, and distraction. *Mind and Language*, 29(1), 26–50.
- Kurtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in processing object shape. *Journal of Neuroscience*, 20(9), 3310–3318.
- Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences*, 3, 254–264.
- Long, G. M., & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: Alternating views of reversible figures. *Psychological Bulletin*, 130(5), 748–768.
- Long, G. M., Toppino, T. C., & Kostenbauder, J. F. (1983). As the cube turns: Evidence for two processes in the perception of a dynamic reversible figure. *Perception and Psychophysics*, 34(1), 29–38.
- Lowe, D. (1984). Perceptual organization and visual recognition. Unpublished doctoral dissertation, Stanford University, Stanford, CA.
- Mathes, B., Strüber, D., Stadler, M. A., & Basar-Eroglu, C. (2006). Voluntary control of Necker cube reversals modulates the EEG delta- and gamma-band response. *Neuroscience Letters*, 402, 145–149.
- McCone, E., Kanwisher, N., & Duchaine, B. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Science*, 11, 8–15.
- Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, 4, 539–551.
- Moreno-Bote, R., Rinzel, J., & Rubin, N. (2007). Noise-induced alternations in an attractor network model of perceptual bistability. *Journal of Neurophysiology*, 98, 1125–1139.
- Mulder, J. A., & Dawson, R. J. M. (1990). Reconstructing polyhedral scenes from single two-dimensional images: The orthogonality hypothesis. In P. K. Patel-Schneider (Ed.), *Proceedings of the 8th Biennial Conference of the CSCSI* (pp. 238–244). Palo Alto, CA: Morgan-Kaufmann.

- Ogawa, Y., Isokawa, T., Matsui, N., Murata, T. (2000). "A neural network model for perceptual alternation of ambiguous figures". Proceedings IEEE Intl. Workshop on robot and human interactive communication, pp. 264–269.
- O'Reilly, R., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience*. Cambridge: MIT Press.
- Peissig, J. J., Wasserman, E. A., Young, M. E., & Biederman, I. (2002). Learning an object from multiple views enhances its recognition in an orthogonal rotational axis in pigeons. *Vision Research*, *42*, 2051–2062.
- Perkins, D. N. (1968). "Cubic corners". Quarterly Progress Report Research Lab in Electronics, No. 89, Massachusetts Institute of Technology.
- Perkins, D. N. (1972). Visual discrimination between rectangular and nonrectangular parallelepipeds. *Perception and Psychophysical*, *12*, 396–400.
- Perkins, D. N. (1971). Geometry and the Perception of Pictures: Three Studies, Project Zero, Technical Report no. 5. Cambridge, MA: Harvard University Press.
- Peterson, M. A., & Gibson, B. S. (1991). Directing spatial attention within an object: altering the functional equivalence of shape descriptions. *Journal of Experimental Psychology, Human Perception and Performance*, *17*(1), 170–182.
- Pheiffer, C. H., Eure, S. B., & Hamilton, C. B. (1956). Reversible figures and eye-movements. *American Journal of Psychology*, *69*, 452–455.
- Pylkkänen, L., & McElree, B. (2007). An MEG study of silent meaning. *Journal of Cognitive Neuroscience*, *19*(11), 1905–1921.
- Reisberg, D., & O'Shaughnessy, M. (1984). Diverting subjects' concentration slows figural reversals. *Perception*, *13*(4), 461–8.
- Richards, J. E., & Gibson, T. L. (1997). Extended visual fixation in young infants: Look distributions, heart rate changes, and attention. *Child Development*, *68*, 1041–1056.
- Rock, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.
- Romei, V., Brodbeck, V., Michel, C., Amedi, A., Pascual-Leone, A., & Thut, G. (2008). Spontaneous fluctuations in posterior alpha-band EEG activity reflect variability in excitability of human visual areas. *Cerebral Cortex*, *18*(9), 2010–2018.
- Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., & Poggio, T. (2007). Robust object recognition with cortex-like mechanisms. *IEEE Transaction Pattern Analysis and Machine Intelligence*, *29*(3), 411–425.
- Seubert, J., Humphreys, G. W., Muller, H. J., & Gramann, K. (2008). Straight after the turn: the role of the parietal lobes in egocentric space processing. *Neurocase*, *14*(2), 204–19.
- Shepard, R. N. (1981). Perceptual organization. In M. Kubovy & J. Pomerantz (Eds.), *Psychophysical complementarity*. Hillsdale, NJ: Erlbaum Associates.
- Shepard, R. N. (1990). *Mind sights*. New York: Freeman and Company.
- Shevelev, I. A., Lazareva, N. A., Novikova, R. V., Tikhomirov, A. S., Sharaev, G. A., & Cuckiridze, D. Y. (2001). Tuning to Y-like figures in the cat striate neurons. *Brain Research Bulletin*, *54*(5), 543–551.
- Strüber, Daniel, & Stadler, Michael. (1999). Differences in top-down influences on the reversal rate of different categories of reversible figures. *Perception*, *28*, 1185–1196.
- Suppes, P., Cohen, M., Laddaga, R., Anliker, J., & Floyd, R. (1983). A procedural theory of eye movements in doing arithmetic. *Journal of Mathematical Psychology*, *27*(4), 341–369.
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, *19*, 109–139.
- Treuhub, A. (1991). *The cognitive brain*. Cambridge: MIT Press.
- Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
- Vidal, J. R., Chaumon, M., O'Reagan, J. K., & Tallon-Baudry, C. (2006). Visual grouping and the focusing of attention induce gamma-band oscillations at different frequencies in human magnetoencephalogram signals. *Journal of Cognitive Neuroscience*, *18*(11), 1850–1862.
- Vogels, R., Biederman, I., Bar, M., & Lorincz, A. (2001). Inferior temporal neurons show greater sensitivity to nonaccidental than metric differences. *Journal of Cognitive Neuroscience*, *13*, 444–453.
- Witkin, A. P., & Tenenbaum, J. M. (1983). On the role of structure in vision. In J. Beck, B. Hoep, & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 481–543). Orlando, FL: Academic Press.
- Ware, C. (1995). Dynamic stereo displays. Proceedings of the SIGCHI conference on human factors in computing systems (pp. 310–316).
- Wright, R. D., & Ward, L. M. (2008). *Orienting of attention*. Oxford: Oxford University Press.

- Wolfe, J. M. (2000). Attention is fast but volition is slow. *Nature*, *406*, 691.
- Wolfe, J. M., & Friedman-Hill, S. R. (1992). The role of symmetry in visual search. *Psychological Science*, *3*(3), 194–198.
- Yonas, A., & Arteberry, M. E. (1994). Infants perceive spatial structure specified by line drawings. *Perception*, *23*, 1427–1435.
- Yonas, A., Cleaves, W. T., & Pettersen, L. (1978). Development of sensitivity to pictorial depth. *Science*, *200*, 77–79.
- Young, A. W., & Deregowski, J. (1981). Learning to see the impossible. *Perception*, *10*, 91–105.
- Zwicky, A. M., & Sadock, J. M. (1975). Ambiguity tests and how to fail them. In L. Kimball (Ed.), *Syntax and semantics* (Vol. 4, pp. 1–36). New York: Academic Press.