Perceiving Function and Category

1. Introduction

This report discusses how we humans perceive functions and categories of objects around us. Such knowledge is extremely critical to our survival even though we often take it for-granted. Just to illustrate the point, imagine being transported to a new planet (say Mars). Human vision system would allow us to see the 3-dimensional structure of all the objects around us and we will be able to navigate our way without colliding into objects. However, questions like '*which objects are edible?*' or *'Which objects are harmful?'* would depend on our ability to perceive functions and categories. Hence, perceiving functions and categories is a very important task of our visual system.

The organization of this report is as follows. In section 2, we discuss the two different frequently described modes (direct and indirect) of human function perception. In section 3, we discuss multiple phenomena which are a result of the way humans perceive categories of objects. In Section 4, we try to describe these phenomena using some widely cited theories. Lastly, we focus on a specific application domain of categorization, i.e. identifying letters and words in section 5.

2. Perception of Function

There are 2 major theoretical approaches to the visual perception of function as shown in Figure 1.

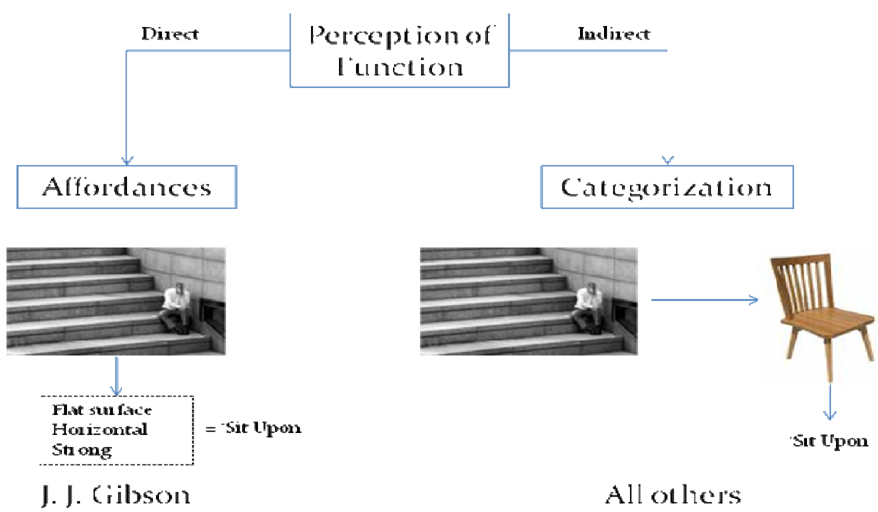


Figure 1: Direct and indirection modes of perception

The *direct* approach states that an object's *opportunities for action (*or *affordances)* can be directly perceived from its visual structure. Hence, according to this theory, once we see 'stairs', our brain straightaway tells us that they can be 'sit upon'. The *indirect* approach states that when we see objects, we internally classify them into known categories with known functions. Hence, according to this theory, once we see 'stairs', we try to match it with the known categories in brain and upon a matching with 'chairs' we decide that it can be 'sit upon'.

2.1 Direct perception

Majority of the work on perceiving functions has been done by the advocates of the indirect theory. However, one noticeable exception is J.J. Gibson (1979). He advocated that whether an object affords to being 'grasped', 'sat upon' or 'used for cutting' can often be decided without its categorization as a 'ball', 'chair', or a 'knife'.

There are two important considerations for Gibsonian affordances:

1. Functional form: This means that an object's function must follow its *form (*i.e. the physical structure). For example, an object must demonstrate that it has the height, stability and 'a level top' for it to be perceived as *'sittable upon'*. An arbitrary shaped structure with no flat surfaces obviously cannot be perceived as 'sittable upon'.
2. Observer relativity: This means that the affordance of an object depends on who is looking at it. For example, a chair may provide 'climbable upon' affordance for a small kid, but would not offer the same for an adult.

These 2 conditions are often together known as 'physical affordance'. Note however that, while these 2 physical affordance conditions are *necessary* for direct perception, they are *not sufficient* for it. Thus objects with the same physical structure may often have different functions. Figure 2 clearly highlights this point.

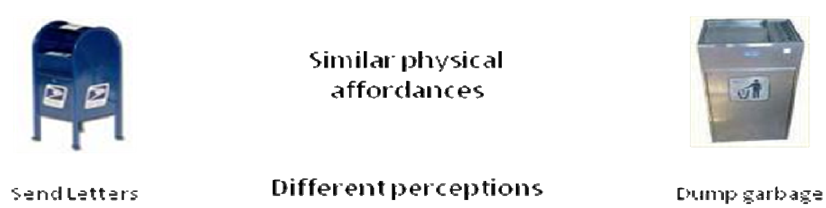


Figure 2: Objects with same physical affordance but different functions.

Clearly, the functions of all objects cannot be perceived directly. For example, it is highly unlikely that a DVD's shape itself would immediately indicate its function. Hence a large amount of function perception is undertaken by object's categorization into a known type.

Neisser (1989) had suggested that affordance and categorization are two fundamentally different modes of perception handled by separate parts of the brain. Medical evidence from patients with damaged ventral systems has confirmed such notions.

2.2 Indirect perception

Indirect perception happens when the object is mapped to a known *category* before its function is defined. There are 4 components in the process of object categorization:

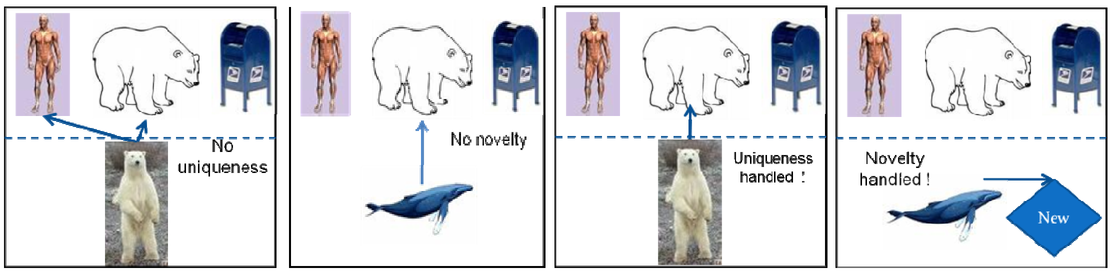
1. *Object representation:* The relevant characteristics of the to-be characterized object need to be perceived and presented within the visual system.
2. *Category representation*: Each of the possible categories should be represented in the memory in a way that is accessible to the visual system.
3. *Comparison process*: There should be a way by which the object can be matched or compared with the possible categories.
4. *Decision process:* This should be a method for deciding the category to which an object belongs, on the basis of the comparison results obtained.

While there is no consensus on how the objects and categories are represented in the brain, it is agreed that their representation should be in a *same* format because the objects need to be compared with the categories. It is largely believed that the comparison process across categories happens 'in parallel' rather than 'in serial'. The intuitive argument is that sheer number of known categories in the brain means that such a process cannot be undertaken in real-time if undertaken serially.

There exist a number of theories about how the 'Decision process' takes place in our brain. Some of the most commonly cited methods are:

1. *Threshold rules:* This method suggests that a criteria value outcome of the comparison process is used and the object is assigned to the category (if any), for which the match exceeds a certain threshold. Its drawback is its inability to handle uniqueness as the object to be categorized may exceed the threshold of more than one category at the same time.
2. *Maximum (Best-fit) rules*: This approach works on selecting the category which gives best-fit amongst all the categories compared. This method cannot handle novelty, as all objects will always get classified into one of the known categories.
3. *Maximum-over-threshold rules:* This approach sets a threshold below which an object is perceived as novel and then employs a maximum-fit if the object is not novel. It obviously can handle both uniqueness and novelty aspects.

This point has been highlighted in Figure 3. A ‘standing bear’ may cross the thresholds for both a ‘standing human’ and ‘bear’ Figures if just thresholds are used. Similarly a ‘dolphin’ may get classified as the nearest matching ‘bear’ Figure if there is no scope to identify new categories as in best-fit rules. Maximum over-thresholds can handle both the type of scenarios.



(a)Threshold rules (b) Maximum (best-fit) rules (c) & (d) Maximum-over-threshold rules

Figure 3: Maximum over threshold scheme can handle both uniqueness and novelty

3. Phenomena of perceptual categorization

In this section we discuss the various phenomena of human visual perception related to categorization.

3.1 Categorical hierarchies

We typically classify objects into categorical hierarchies (e.g. Doberman<Dog < Mammal < Animal < Living thing…). The classical or *Aristotlean* view on categorization was that there exist a set of necessary and sufficient conditions which classify objects into categories. Hence an object exists in the binary state of either ‘belonging to’ or ‘not belonging to’ a particular category. However, there exist multiple instances where there are no set of rules can define a category. For example, defining such a set of rules for the category of ‘games’ is extremely difficult. Hence, Wittgenstein (1953) argued that categories cannot have a binary classification.

The most prominent theory on categorization was proposed by Rosch in 1970s and is based on ‘*prototypes’.* A prototype is the ‘average’ member of a class which can be used to represent it. Hence a ‘doggiest possible dog’, would form a ‘prototype’, for the dog class. Thus ‘typicality’ turns out to be an important factor for object classification.

An object typically gets classified at multiple levels (e.g. Lassie < Doberman < Dog < Mammal < Animal). Most people however, tend to classify objects at a certain intermediary level (e.g. dog) known as the *‘basic level category’*. The levels above basic are known as the super-ordinate categories and those below are known as the subordinate categories. Work by Rosch and others suggests that basic level category is the highest category level at which the objects have similar shapes, similar motor interactions and common attributes.

Work by Jolicoeur (1984) suggests that 'typical' members tend to be classified at the 'basic' level, however 'atypical' members tend to be classified at subordinate levels. For example, the first word which comes in mind of people for 'Robin' (Figure 4 (a)) is quite likely to be 'bird', but that for a 'Penguin' (Figure 4 (b)) is quite likely to be 'Penguin' itself. Jolicleur called such subordinate categories which are preferred by humans as the *'entry level categories'*.

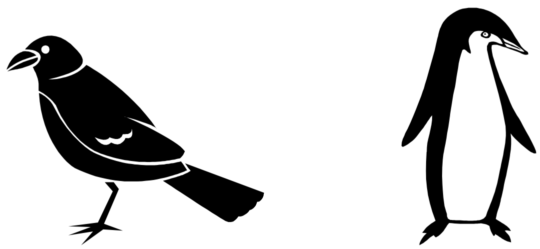


Figure 4: (a) 'Typical' and (b) 'Atypical' members of bird class.

3.2 Perspective viewing

The perspective from which we view an object also influences the speed and accuracy of recognition. For example, as shown in Figure 5, a horse can be identified much faster and accurately from a side-view than a top-view.

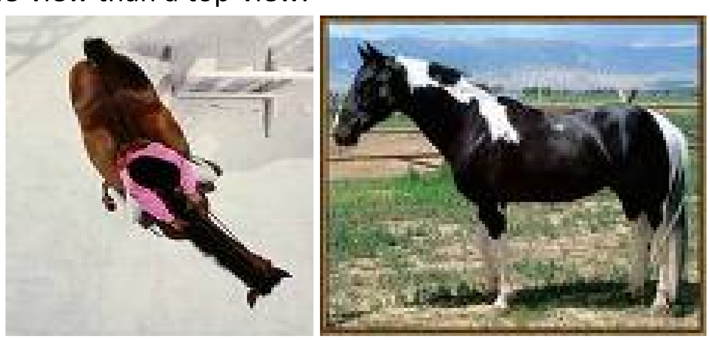


Figure 5: A horse seen from two different perspectives.

The best, most easily identifiable view for each object is known as the *'canonical perspective'.* Canonical perspectives are often explained in terms of 'frequency hypothesis' (i.e. how often we see the object from that viewpoint) and 'Maximal information hypothesis' (i.e. best views tend to show multiple sides of objects)

Beiderman(1991) studied the impact of perspective on the '*priming effect'* (i.e. objects are recognized faster if shown once earlier). He concluded that priming effect remains intact even with changes in size, position and reflection. However priming effect does get diminished if objects are presented second time from a perspective such that the parts visible are not the same as the first time.

Jolicoeur(1985) studied effect of orientation on the recognition of objects. He noticed that the subjects were fastest at naming objects in their upright orientation. He also noticed that naming latencies increase with the angular distance from the upright position. However, he noticed a significant decrease in delay for 180 degree rotations than what was predicted by the angular difference logic. The overall effect is shown in Figure 6.

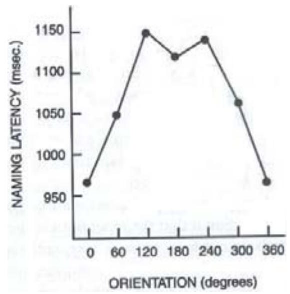


Figure 6: Impact of orientation change on naming latency

3.3 Part structure

An important point studied by researchers is whether the object categorization is influenced by parts of the object which are presented to the subject.

Bederman et al.(1991) studied the impact of the part structure using the priming effect concept. They noticed significant priming effects even when half the component lines of an image were removed, as the part structure was still obvious. On the other hand when they changed the part structure, the priming effect was significantly reduced.

3.4 Context

Categorization of an object also gets influenced by the spatial array of objects surrounding it. This is known as the contextual effect. For example in Figure 7, the same shape is once perceived as 'H' and once as 'A', due to the surrounding objects. Similarly, a group of fruits get classified as a face just due to their spatial correlation. Studies by Palmer(1975) have confirmed that appropriate context facilitates categorization, while inappropriate context may hinder it.

Figure 7: The effect of context in perception

3.5 Visual agnosia

A distinct type of phenomena regarding categorization studied often from a biological perspective is that of 'visual agnosia'. Visual agnosia refers to the perceptual deficit due to damage to the brain, in which patients are unable to correctly categorize objects they were previously familiar with.

Apperceptive agnosia refers to the inability to perceive categories due to damage to the sensor system. *Associative agnosia* refers to the state where patients with fully working perceptual abilities are unable to categorize the objects they see. An example of such phenomena is shown in Figure 8, which shows that a patient could visually see and replicate drawings of common objects but was unable to name them.

Another classic example of visual agnosia is the phenomena of *Prosopagnosia* which refers to the inability to recognize faces visually. Qoute from Pallis (1955), reveals the seriousness of the phenomena. "At the club I saw someone strange staring at me, and asked the steward who it was. You'll laugh at me. I'd been looking at myself in a mirror."

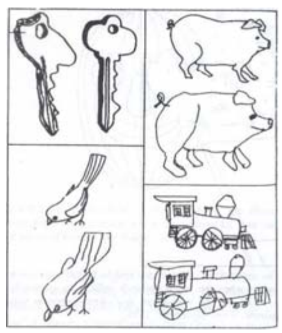


Figure 8: Impact of Associative agnosia

4. Theories of object categorization

Having discussed multiple categorization phenomena which are demonstrated by the human visual perception, now we study some generic theories which try to explain them. We will start by describing Irving Biederman's 'Recognition By Components' theory and see how well it can explain the various observed phenomena described in the previous section. We will then do a quick study of few other alternate theories proposed in literature and compare them with RBC.

4.1 Recognition By Components (RBC) theory

This theory works on the premise that objects can be specified as spatial arrangements of primitive volumetric components known as *'geons'.* The term 'geons' means 'geometric ions' and thus indicates that these are the building blocks of various objects. The relation between geons and objects can be understood to be similar to that between 'words' and their constituent 'letters'.

Another way to look at geons is as a set of generalized cylinders which are easily distinguishable from each other. This is illustrated in Figure 9. In the Figure shape A is used as a standard, and the variations in terms of different features are shown by its comparison with other shown shapes (B, C, D …). The properties on which the geons can vary are:

1. *Cross sectional curvature*: which can be straight (A) or curved (B)
2. *Symmetry*: The symmetry across the axis can be none (D), reflectional (C) or both (A)
3. *Axis curvature*: The sweeping axis can be straight (A) or curved (E)
4. *Cross sectional size variation*: The size of the cross sections can be constant (A), expanding (G) or expanding and contracting (F).
5. *Aspect Ratio*: The ratio of the length of the sweeping axis to the cross section can be approximately equal (A), axis greater (H) or cross-section greater (I).

Out of these 5 properties, the first 4 are considered to be 'qualitative' and the last one is considered 'quantitative'. Hence by considering the possible combinations across each of the properties and the number of possible values, we can find that there are 36 (2 X 3 X 2 X 3) qualitative different geons. The number goes up to 108 (3 X 36), if we also consider the 'quantitative' property of Aspect Ratio in variation.

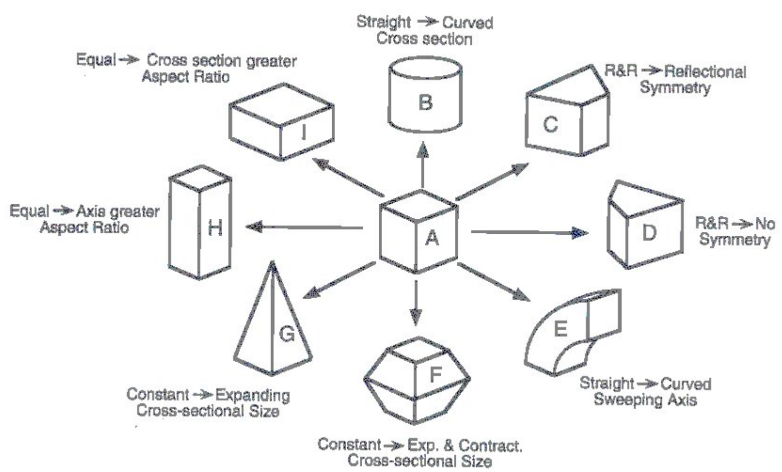


Figure 9: Different types of geons

In order to create shapes which are complex than one geon, their *spatial relationships* are used. The use of spatial relationship between geons is similar to the idea of using order of letters to create different words in the aforementioned analogy. An example of the use of same geons to create different objects is shown in Figure 10.

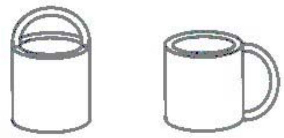


Figure 10: Same geons with different relations

Geon relations concern how they are attached (Side-Connected, Top-Connected etc.) as well as their relational properties like relative size (Larger-than, Smaller-than etc.). In all there are 108 different geon relationships defined. It might be interesting to note that these 108 different relations can lead to more than 1 million different shapes for just 2 geon objects.

A neural network implementation of the RBC was undertaken in 1992 by Hummel and Biederman. An overview of the implementation is shown in Figure 11, which describes how the layers would work to detect a simple 2 geon object shown at the bottom of the Figure.

The recognition process starts at layer 1, where the image edges are detected. This is followed by layer 2 where low level image features like Vertices, Axes and Blobs are detected. Layer 3 identifies the correct geon attributes for the two physical geons (shown in yellow) as well as the relationship between them (shown in red).

These geon relationship attributes are combined into actual relations in layers 4&5. Layer 6 uses the geon attributes found in layer 3 together with the relations found (layers 4&5) to identify the correct geons present. The combination of geons finally leads to object recognition in layer 7.

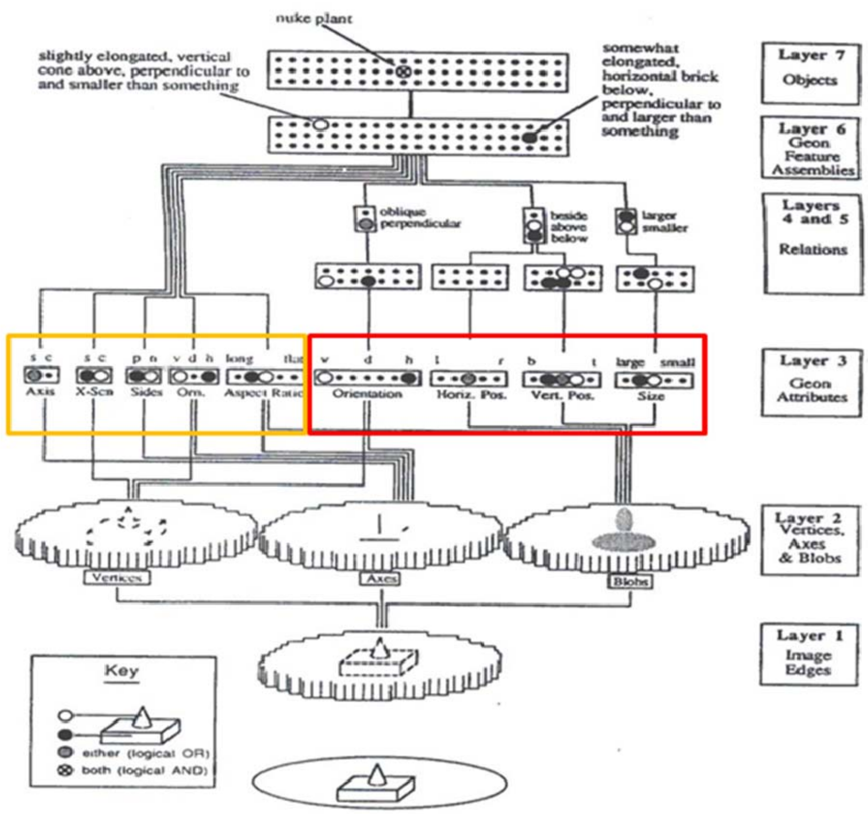


Figure 11: A neural network implementation of RBC theory

4.2 Explaining the observed phenomena

Now let us see how well the RBC theory can explain the various phenomena of visual perception that we discussed in the previous section.

1. *Categorical hierarchies*

   The notion of *prototypes* and *typicality* can be very well explained by the geon theory. Our brain contains a mental model for the prototype of each category. Clearly this model which uses geons will be at a qualitative level and cannot capture the fine grain details.

   Thus it can be argued that geon matching results in object classification and the typical members usually do get classified correctly at the basic category level as they match the mental geon model. The atypical members do not match the mental geon model and hence need to be classified at a subordinate category. Figure 12 (a) shows a possible geon based representation of a 'bird' in the brain. Hence a 'robin' (b) is likely to be classified at the basic level, while a 'penguin' is unlikely.

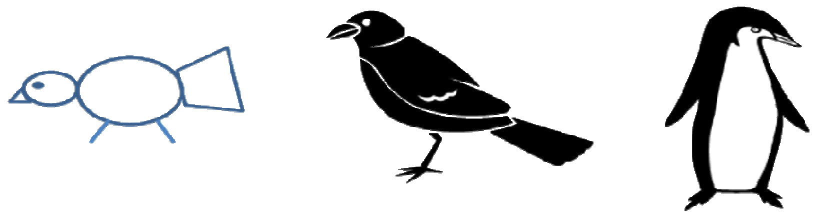


Figure 12: (a) A mental geon-based model of a 'bird' (b) a typical member (c) an atypical member

2. *Perspective viewing*

   The *canonical perspective* phenomena can also be explained using the geon theory. If we consider the same notion of a representative geon model in our brain, the perspective from which the image obtained best matches with the mental model is indeed the canonical perspective. Other perspectives may not be as useful for categorization as some of the geons get occluded in those views. This point is highlighted in Figure 13, as the mental model (Figure 13(a)) shows the same set of geons as the canonical perspective for the horse category (Figure 13(b)). The top-view does not show the same set of geons and hence is difficult to identify or categorize.

Figure 13: (a) A mental geon-based model of a 'horse' (b) canonical perspective (c) top-view

3. *Part structure*

   The part structure phenomenon is also well explained by the geon theory. As it is argued that the human categorization works on geons rather than specific lines or contours, it makes sense for the priming effect to be visible irrespective which component lines are shown to represent that geon. For example, the 'Identity' and 'Line Complement' images shown in Figure 14(a), both create the same geons in the brain and hence create similar priming effects for a 'piano' object (Figure 14(b)). However, the 'Different exemplar' image shows different parts and geons, and hence shows reduced priming effect.

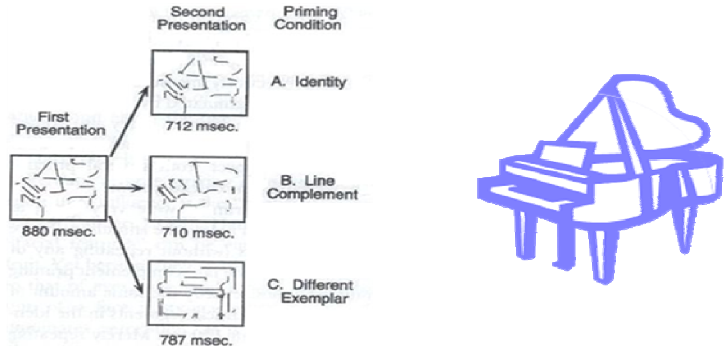


Figure 14: (a) Priming effect between first and second presentation of an object under different conditions (b) the 'Piano' object

4. *Context*

   The contextual effects cannot be explained by RBC as it focuses on parts *within* an object rather than those surrounding it. However, there has been another set of research (Palmer, 1975), which classifies the entire scene as an 'object' and tries to explain context as an interaction between its parts.

5. *Visual Agnosia*

   One reasonable explanation for *'Associative agnosia'* is that the brain loses its ability to store mental models of categorical prototypes. Hence all views of objects are totally new views to it and thus cannot be used for categorization.

   However, it must be noted that it is quite difficult to generically explain the visual agnosia like phenomena with the RBC theory, because these phenomena happen inside the brain and there are no easy ways to understand what is happening there.

4.3 Other theories

As is clear from our discussion, while RBC can explain many of the categorization phenomena, it cannot explain all of them. Its prowess is limited by the (only) 108 different type of geons which are not sufficient to capture (fine-grained) differences often required to say distinguish between different families of cats.

Further, it has been argued in literature that humans do not store just one view of an object in the brain as that would definitely not suffice for 3D object recognition. This has resulted in multiple viewpoint specific theories. We take a quick look at them.

1. *Aspect Graphs*

   Aspect graphs work on the idea that many views of the same object are actually very similar. This common abstract representation of views is known as an *'aspect'*. This has been shown in Figure 15. In Figure 15(a), while the first two views of the presented object are slightly different, they are classified as one aspect because they show the same set of sides and edges and hence can be given a single abstract representation. The 3rd view on the other hand shows different sides and edges and hence is labeled as a separate aspect. Similar representations for a 'tetrahedron' have been shown in Figure 15(b). The common representation shown at the bottom captures the possible linkages between the edges in an abstract manner.

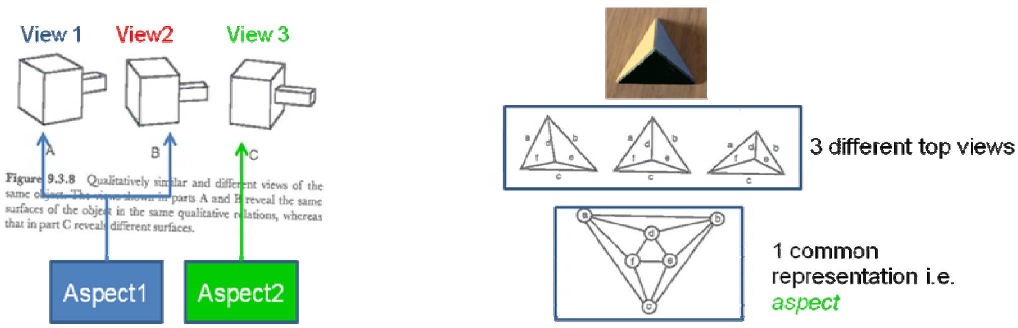


Figure 15: Different views resulting in common aspects

The 14 possible topologically distinct views or aspects of a tetrahedron are shown linked with edges in Figure 16 (Different faces are indicated by different shading) to create an *Aspect graph*. Two stable views are marked as *adjacent,* when it is possible for a moving viewer to pass from one view to the other with no intervening stable views. In the aspect graph, each pair of nodes representing adjacent stable views is joined by an edge. (Koenderink & van Dorn, 1979 and Schiffenbauer,2001)

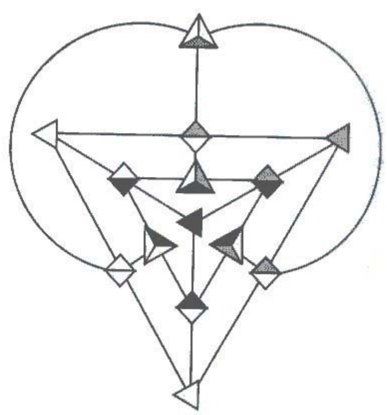


Figure 16: A network connecting different aspects of an object

While Aspect graphs successfully provide abstract representation for the objects, they also suffer from multiple issues. Firstly, they do not scale well to complex shapes. Secondly, they capture only the topology and not the geometry of objects. Hence they cannot differentiate between a cube, cuboid and an irregular parallelepiped.

2. Alignment with 3-D models

Lowe (1985) and Ullman (1989) have argued that models of all known objects are stored in the brain in a 3D representation. To categorize an object, first its salient edges are detected. Then these salient edges are matched with the 3-D models present in the brain, and the object is finally recognized if there is a close match found. The problem with this theory is that it assumes that 3D models of all objects are already present in the brain.

3. Alignment with 2-D view combinations

A slightly different notion is presented in works by Poggio(1990) and Ullman(1996). They argue that the brain has multiple 2D representations of objects rather than a 3-D model. Newer 2D representations can be obtained by combination of known 2D representations. While this theory explains well the presence of multiple 'good' views for recognition, it also assumes that representations of all objects are already present in the brain.

5. Identifying letters and words

Identifying letters and words can be considered a specific application of object categorization which is frequently encountered in practice. In this section we will discuss visual perception works for identifying letters in isolation and also when they are present as parts of words. We will then try to explain the processes using the widely studied Interactive Interaction Model and also glance over some of the other possible explanations in literature.

5.1 Identifying letters

Letters can be identified by using the standard methods for 2-D object identification such as Templates, Structural descriptions or Features. Features are indeed an interesting possibility and work by McCelleland and Rumelhart(1981), describes how the English uppercase letters can be represented using features like number of vertical lines, horizontal lines, L-junctions, T-junctions etc. such that no two letters have the same features.

5.2 Identifying letters within words

Psychological analysis has shown that letter identification is significantly influenced by the word that it is part of. A simple but powerful example of this has already been shown in 'THE CAT' example shown in Figure 7. As another example, if we gave viewers 5 seconds to look at the first line of text in Figure 17 and asked them how many letters they can remember, and then repeated the same experiment with the bottom line, most people will perform much better with the bottom line.

HWO NMYA RSETELTE NCA OYU RPTERO WNO

HOW MANY LETTERS CAN YOU REPORT NOW?

Figure 17: Words can influence the letter identification

This phenomena that the words can influence how we identify letters is known as the *'word superiority effect'* and was first reported by Cattell (in 1886!). This phenomena was re-examined by Reicher (1969) and Wheeler (1970), who worked towards separating out the effects of *remembering* from *identification* of letters.

A controlled experiment as shown in Figure 18, however again demonstrated the word superiority effect. In this experiment, the subjects were shown a 'Target Display' for 50 ms, followed by a masking pattern over the area that contained the crucial letters. Just above the mask, the users were presented with two choices 'D' and 'K' and asked to tell which letter had been present at that position. As the numbers on the right indicate, the accuracy for correct letter identification was the highest when letters formed part of a valid English word. The rate reduced significantly when a non-word anagram was presented or even when just a single letter was presented.

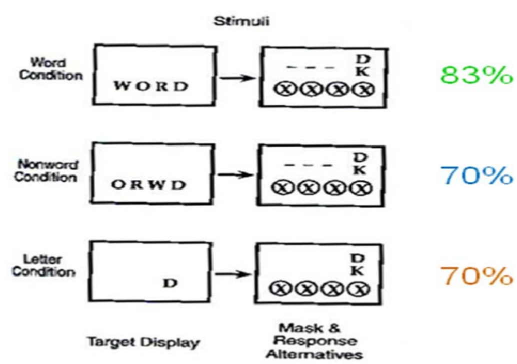


Figure 18: Word superiority effect

The phenomenon that letters are more readily identified in words than random strings is known as the *word-nonword effect* and the phenomena that a single letter is perceived less accurately than when it is part of a word is known as the *word-letter effect.*

5.3 Interactive activation model

Interactive Activation Model is a theory proposed by McClelland and Rumehart (1981), which tries to explain the phenomena observed in letter and word identification. As shown in Figure 19, this theory suggests 3 components of the letter/word identification process.

The visual input is processed at the feature level to help in letter identification. The letter identification step in turn helps with the word identification. The interesting aspect however is that the word level identification component also provides *bias* as to which letters can be present to the letter level component based on a dictionary of valid English words. The interactive activation model suggests that such a process takes via neural network like procedure where the feature level nodes excite possible letters and the word level nodes excite or inhibit the various letter nodes.

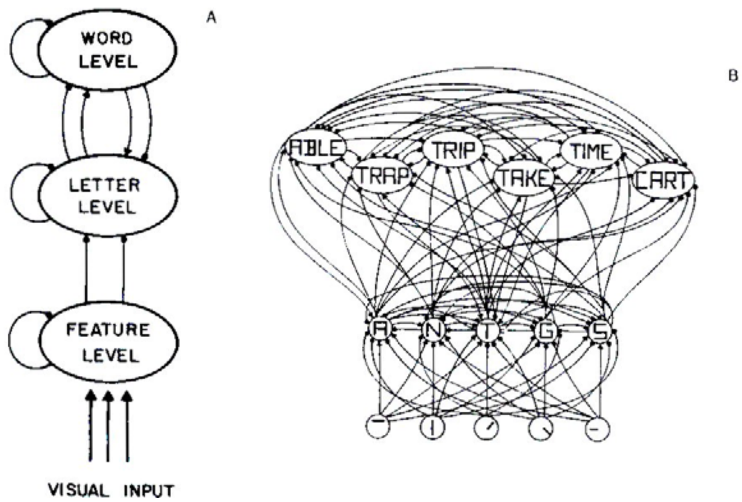


Figure 19: Interactive Activation Model

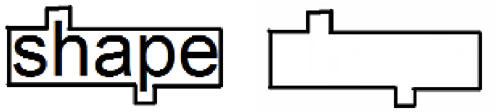
This model can explain the 'word superiority' phenomena as those letters which are part of valid English words get excited more often. Similarly, word-letter effect is explained as single letters get excitation only from the feature level nodes and not the word level nodes.

5.4 Alternative explanations

Related research also points to some very interesting alternative explanations about how words and letters are identified.

1. Word Shape

James Cattell (yes 1886!) suggested that we identify words as a pattern of ascending and descending *shapes*. Modern experiments conducted independently seem to corroborate his theory. For example, as shown in Figure 20 (c), if the word 'test' was miss-spelt and multiple subjects were asked to proof-read the text containing it, it was found that the error went unnoticed more often if the changed word had the same shape as the original word (e.g. 'tesf' for 'test') than otherwise.



test	Error rates	Explanation
tesf	13%	Consistent word shape
tesc	7%	Inconsistent word shape

Figure 20: (a) A word with its shape, (b) just the shape and (c) Word shape used for word identification

2. Serial letter identification

Gough(1972) advocated that the letters contained in words are processed serially from left to right. He provided an analogy to the dictionary and pointed out that bigger words take longer to identify. An example of such serial effects that the 1st letter is processed before 2nd and so on is shown in Figure 21. Most people will read the 1st line as North and the 2nd line as South, even though exactly one letter needs to be changed to read it either ways.

NOUTH

SORTH

Figure 21: Serial letter identification effect

The serial letter identification theory however fails to explain the word superiority effect.

3. Parallel letter identification

The majority of research community argues for the parallel letter identification process. McConkie and Rayner (1975) describe a moving window effect which says that humans jump between points in text at which they *fixate* as shown in Figure 22. The perception system typically spends 200-300 ms to fixate on words followed by 20-35 ms of *saccadic movement.* While fixating, 3 to 4 letters on both sides of the fixation point are actually read in the fovea and around 15 to 20 characters (on both sides) are perceived at the parafovea to influence about the context as well as decide on the next fixation point.

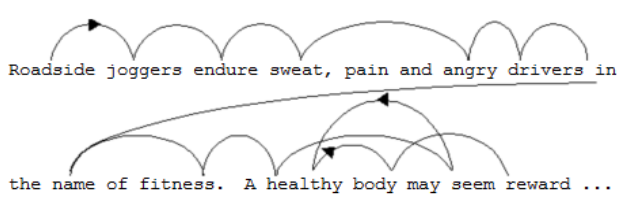


Figure 22: Saccadic movements and fixation points in reading

Such notions have been verified by experiments in which part of the text beyond the fixation point was masked and revealed only *during* the time when the eye undertook saccadic movement to move to the next fixation point. The experiments showed that the reading speed increased as the number of characters visible to the left of fixation point was increased from 3 characters to 15.

Rayner (1975) also conducted similar experiments in which he did not mask the text but rather showed different text which was controlled to have different features than the actual word. As summarized in Figure 23, the time taken to read the words was fastest when the correct word was allowed to be read in the parafovea. It decreased successively as the word presented were increasingly dissimilar in terms of shape and letters used.

Word shown	Properties	Speed
chart	Identical word (control)	210ms
chovt	Similar word shape Some letters in common	240ms
chyft	Dissimilar word shape Some letters in common	280ms
ebovf	Similar word shape No letters in common	300ms

Figure 23: Different words shown in para-fovea and its impact on reading speed

Suggestions for further reading

Palmer1999: Stephen E. Palmer, Vision Science: Photons to Phenomenology, The MIT Press. ISBN-13: 978-0-262-16183-1. (*Primary source*)

Gibson1979: Gibson, J.J., The Ecological Approach to Visual Perception. Boston: Houghton Mifflin.ISBN 0898599598, 1979.

Mervis1981: Mervis, C. B., & Rosch, E. Categorization of natural objects. In M. R. Rosenzweig & L. W. Porter (Eds.), Annual Review of Psychology (Vol. 32). (1981).

Biederman1987: I. Biederman. *Recognition-by-components*: A theory of human image understanding. Psychological Review, 94(2):115--147, 1987.

Hummel1992: Hummel, J. E. & Biederman, I. Dynamic binding in a neural network for shape recognition. Psychological Review, 99, 480-517. (1992).

Ullman1996: High Level Vision, by Shimon Ullman, MIT Press (Bradford), Cambridge, Mass., xviii+412 pp, ISBN 0-262-21013-4. (1996)

Cattell1886: Cattell, J. The time taken up by cerebral operations. *Mind*, 11, 277-282, 524-538. (1886).

Gough1972: Gough, P.B. One second of reading. In Kavanagh & Mattingly's *Language by ear and by eye*. Cambridge, MA: MIT Press. (1972).

McConkie1975: McConkie, G.W. & Rayner, K. The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics*, 17, 578-586. (1975)

Rayner1975: Rayner, K. The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 7, 65-81. (1975).

Schiffenbauer2001: Robert D. Schiffenbauer, A Survey of Aspect Graphs, Technical report-Polytechnic University, 2001.

Reading_Game: *Remove 50% of the letters from the text and you can still read the text*! http://www.ababasoft.com/words/remove_letters.html

AspectGraphs: Tutorial. http://people.csail.mit.edu/bmcutler/6.838/project/aspect_graph.html