

Bachelor Thesis

The Disjunction Problem and Asymmetric Dependency

Does Fodor solve the problem?

Supervision: Menno Lievers

Jeroen Witteveen

3362264

30 August 2012

Bachelor of Philosophy

Utrecht University

2008 – 2012

Table of contents

Introduction	3
1. Computational Representational Theory of Thought	4
1.1 Computations.....	4
1.1.1 <i>Propositional attitudes and the language of thought</i>	4
1.1.2 <i>Computations</i>	5
1.1.2 <i>The computational component</i>	6
1.2 Representations.....	6
1.2.1 <i>The external approach</i>	7
1.2.2 <i>The Causal Theory of Perception</i>	7
1.2.3 <i>The representational component</i>	8
1.3 Final remarks	8
2. Problems in the Causal Theory of Perception	9
2.1 The disjunction problem	9
2.2 The Asymmetric Dependency Theory.....	10
2.3 Problems with the Asymmetric Dependency Theory	10
2.3.1 <i>Names</i>	11
2.3.2 <i>Non-existing objects</i>	12
2.3.3 <i>Lookalikes</i>	12
2.4 Final remarks	14
3. The deeper problem	15
3.1 Normativity in perception	15
3.2 Solution?.....	15
3.2.1 <i>Abortive actions</i>	16
3.3 Final remarks	17
4. Conclusion	18
Literature	19

Introduction

The Computational Representational Theory of Thought (CRTT)¹ is a theory that tries to explain certain features of the mind, for example the process of thinking or how there can be meaning in the mind. The idea that the mind somehow works like a computer is often used by adherents of the CRTT to explain these features. However, as Georges Rey points out, “characterizing the brain as a computer with semantic states is an explanatory proposal freed from the conventions surrounding artifacts.”² So CRTT does not say that the mind *is* a computer, it says that the computer is only used *as a model* to characterize a non-artefact natural object, the mind. The mission of CRTT could be defined as follows: “According to CRTT, the mind (and/or brain) is to be understood as a computer with a very specific structure, or ‘computational architecture.’ Specifying that architecture in adequate detail is the burden of CRTT.”³ So CRTT is supposed to explain the computational architecture of the mind.

CRTT can be divided into two components. The first is the computational component, which is about the structure of our thinking. The second is the representational component, which is about how there can be meaning in the mind. To put it bluntly, the computational part refers to the syntax (form) and the representational part refers to the semantics (content) of our thoughts.

My aim is to investigate the representational component, because there are some serious problems with the CRTT’s view of how our thoughts about the world around us can be meaningful. The idea is that, if an object causes a representation via a perceptual process, then that object is the meaning of the representation. However, establishing the connection between a representation and an object in order to ground meaning is more problematic than it seems and raises considerable issues for a proper CRTT. I will investigate the solutions that Fodor provides for this problem.

I will start by scrutinizing the CRTT to make the problems I want to focus on more understandable. Then I will examine how a causal theory of perception should look like to bring semantics to the mind. A central problem with a causal theory of perception is the disjunction problem, which I will elaborate on. Next I will discuss the Asymmetric Dependency Theory as a solution to the disjunction problem as proposed by Fodor and some limitations of this solution. Subsequently, I will focus on an underlying and fundamental problem, namely the problem of normativity in perception, which in short says that a core problem of causal theories of perception is that they can’t explain the connection between an object and a representation. With regard to this problem I will also examine a solution of Fodor, and it will become clear that his solution also falls short as a response to the charge of circularity.

At the end I will answer the question: Does Fodor give a sufficient solution to the disjunction problem and, more fundamentally, to the problem of normativity in the causal theory of perception?

Just a short note before I continue: I will use single parentheses ‘...’ to indicate the mentioning of a word as opposed to using a word without parentheses to stand for an object. I will use double parentheses “...” to indicate mental content and for quoting text. But in the latter case I will add a footnote and also context should make the difference clear.

¹ ‘CRTT’ is a term used by Georges Rey, but the theory is also known as the Representational Theory of the Mind (RTM) or the Computational Theory of the Mind (CTM).

² Rey, G. *Contemporary Philosophy of Mind*, p. 238.

³ *Ibid.*, p. 209.

1. Computational Representational Theory of Thought

The computational representational theory of thought (CRTT) is a theory about the mind that tries to explain certain features of the mind like thinking, intentionality, causality, meaning etc. In particular it is a mentalistic view of the mind, which refers to the idea that physical objects can possess mental properties. Furthermore, it is naturalistic because it assumes that these physical objects with mental properties could be investigated by the natural sciences without assuming the presence of immaterial properties.

As mentioned before, CRTT could be divided in two components: “a computational component that focuses upon computations defined over syntactically specified representations in a specific language, and a representational (or meaning, or “content”) component, that focuses upon how those syntactically specified representations do in fact have specific meanings.”⁴ To separate CRTT in those two components serves mostly an explanatory purpose. I will use this characterisation as the definition of CRTT.

In the next sections I will elaborate on Rey’s definition of CRTT by explaining first the computational component and then the representational component.

1.1 Computations

The first part of the definition of CRTT, given by Rey, says that the “computational component [...] focuses upon computations defined over syntactically specified representations in a specific language.” So to know what the computational component is, we need to know what ‘computations’, ‘syntactically specified representations’ and a ‘specific language’ mean. I will start by explaining what thinking consists of according to CRTT and from there I will try to explain these three concepts.

1.1.1 Propositional attitudes and the language of thought

The CRTT of mind assumes that thinking⁵ requires propositional attitudes. A propositional attitude consists of a proposition, (e.g. snow is white) and an attitude (e.g. believing, hoping, desiring, etc.) towards this proposition entertained by the thinking subject. An example of a propositional attitude is the thought ‘I know that snow is white’. Propositions are also known as ‘content’ and propositional attitudes with content belong to the larger category of ‘mental states’, that also includes perceptual states, emotional states, pain and thoughts directed at individuals as ‘I think about London’. A distinctive feature of CRTT is that it maintains that propositions are sentences in the ‘language of thought’ (LOT). The LOT is the ‘specific language’ in the definition of the computational part of CRTT above. So a propositional attitude is an attitude towards a sentence in the LOT.

Another property of propositional attitudes is intentionality, also called ‘aboutness’. This property means that propositional attitudes are about something, e.g. the belief that snow is white is about that state of affairs that snow is white. So the content of a propositional attitude is

⁴ Ibid., p. 208.

⁵ Whether thinking (or thoughts) is conscious or unconscious is not essential to CRTT. According to Crane we can have unconscious thoughts, which are thoughts that we are unaware of. (Crane, p. 27.) Rey agrees with Crane: “it is enough to see that there is no reason to claim that consciousness is essential to the kinds explanatory mental states captured by CRTT.” (Rey, p. 258.) However there are opponents, like Searle or Strawson, to the idea of unconscious thoughts or mental states. (Rey, p. 256-8) I will not further engage in a discussion on consciousness, because it is not essential in understanding CRTT.

directed at something. This is called *intentionality*. A problem with the notion of intentionality is that it is hard to specify what the real connection between an object and the thought about the object is. Another issue with intentionality is that it seems hard to explain how thoughts can be about non-existing objects like unicorns. The question here is: to what is the intentional state directed if the object does not exist? The problem is that intentionality is indifferent to reality. Hereinafter, I will discuss these issues in more depth.

Another problem with propositional attitudes is their proper interpretation the belief that the sky is blue can be read expressing the belief that there is a thing (*res*), namely 'sky', which has the property of being blue. This is a 'de re' interpretation. However, it could also be interpreted as expressing that I believe that the proposition 'the sky is blue' is true. This is the 'de dicto' interpretation of propositional attitudes. So one problem with respect to propositional attitudes is which of these interpretations, the *de re* or the *dicto* has priority in the order of explanation.⁶ However, this issue falls beyond the scope of this thesis so I will not pursue it here.

The LOT is originally an idea from Fodor. He argues that the LOT is innate, which means that everyone is born with it, and that is the internal language in which we think. All the concepts are already in the head, but they need to be activated, or triggered, by the right perceptual process. So the concepts in the LOT are not learned by experience, but are triggered by it. This provided the hypothesis with a strong explanatory power

However, the idea of the innateness of the language of thought is controversial. Putnam's argument against it is that "evolution would have had to be able to anticipate all the contingencies of future physical and cultural environments. Obviously it didn't and couldn't do this."⁷ I only wanted to point out that the LOT is a controversial notion, so I will also not pursue this issue here.

The sentences of LOT, also known as 'mentalese', are "entokened in the brain, and are subject to certain causal processes."⁸ 'Entokened in the brain' means that they are actual physical entities in the brain that constitute these sentences. Thereby, these sentences are not literally sentences as in a book, but they are "highly abstract objects"⁹, or in other words, a sequencing of tokens in the brain. I will keep using the word 'sentence' to refer to the sequencing of tokens. So if I believe that 'snow is white', then, according to CRTT, there are actual physical entities, for example neuron firings or another physical process, that composes the sentence 'snow is white'.

The mental sentence 'snow is white' is a representation, also called a 'symbol', and it consists of a syntactic and a semantic component. The syntax of the representation 'snow is white' is the form of the representation and the semantics, or meaning, of the representation is that it, in this example, refers to the physical object snow and the property of 'being white'. So the phrase 'syntactically specified representations' only refers to the syntactic properties of the representations, and ignores the semantics of the representations.

1.1.2 Computations

Above I have explained that '*syntactically specified representations over a specific language*' is about the syntactic properties of the symbols in the LOT. These syntactic properties are somehow computed. I will now explain how that works.

⁶ Quine, W. V. 'Quantifiers and Propositional Attitudes' and McKay, Thomas and Nelson, Michael, "Propositional Attitude Reports", The Stanford Encyclopedia of Philosophy.

⁷ Putnam, H. *Representation and Reality*, p. 15.

⁸ Rey, G. *Contemporary Philosophy of Mind*, p. 210.

⁹ *Ibid.*, p. 211.

The syntax in the mind is processed in such a way that a meaningful input leads to a meaningful output. This process can be seen as a computational process. In a computational process a function is computed, which is the same as evaluating a function “by application of a rote procedure.”¹⁰ Take for example the deduction:

$P \rightarrow Q$
P
Thus, Q

This deduction is only based on syntax, there are no semantic notions needed. Whatever P and Q might mean, if we have an input ‘ $P \rightarrow Q$ ’ and ‘P’, then we can automatically conclude ‘Q’ (the output). This process follows the rules of logic and is a computational process, because it follows a ‘rote procedure’. Important is that the truth of the semantics of the symbols are preserved in the process while only syntax is used. If the premises are true, then the conclusion is true as well.

According to CRTT, our mind works according the same principle. We have an input that causes representations, which consists of syntax and semantics. Based on the syntax of the representations, our mind computes according to internal rules and it generates an output in the form of representations, which also consists of syntax and semantics. The internal rules are formed by the structure of our brain and the representations are sentences in the LOT. The process of computing can be seen as neuron firing or other neurological activity and is also called thinking.

1.1.2 The computational component

To repeat: *syntactically specified representations* in a *specific language* are the syntactic properties of the sentences in the LOT. According to the definition of the computational part, ‘computations are defined over’ these syntactic properties. So the syntactic properties of the LOT are computed and this process is called ‘thinking’. An example of thinking is the thought: ‘I believe that p if q’. And if q is the case, then it follows logically that I believe p. The relation between a proposition, a sentence in LOT, and the attitude is called a ‘computational relation’.

There is one more thing to consider about the LOT. The LOT is the language in which we think, it is the content of our thoughts, but it should be distinguished from the meta-language that consists of the ‘rules’ according to which the LOT is computed.¹¹

To summarise this part about computations: the computational component of the CRTT refers to the idea that we think in a language, the LOT, and that thoughts are computed by a set of syntactic ‘rules’, or a meta-language, by virtue of the syntax of the LOT. These thoughts are, according to CRTT, in the form of propositional attitudes.

1.2 Representations

The first part of the definition of CRTT was about the mechanism of thinking. In this section I will consider the second part of the definition of CRTT: “ [...] a representational (or meaning, or “content”) component, that focuses upon how those syntactically specified representations do in fact have specific meanings.”¹² So the question arises: how can syntactically specified representations have specific meanings?

¹⁰ Stanford CTM paragraaf 1.1

¹¹ Ibid., p. 211.

¹² Ibid., p. 208.

1.2.1 The external approach

An account of meaning in the context of CRTT should be realistic, because CRTT treats propositional attitudes as real, physical, entities. Therefore, a notion of meaning should be something 'real' and not just an idealisation or abstraction. Rey considers an internalistic and an externalistic approach to construct a theory of meaning. The internalistic approach assumes that the meaning of a term is determined in the mind by using other internal concepts, however this idea does not suffice according to Rey. He argues that Putnam's Twin-Earth thought experiment¹³ shows us that at least some environmental aspects of the agent are used in determining the meaning of a term.¹⁴ I want to look into the problems of the external approach.

If Rey is right in using Putnam's argument to show that at least some of the environment, or objects outside us, determine meaning, then how do we obtain knowledge about the objects and how can our thoughts have meaning? It seems that there needs to be a connection between the objects and our thoughts in order to establish some notion of meaning, which tells us how a thought can mean a specific object. So knowing how the connection between the object and its representation works is a necessary condition to provide an account of meaning.¹⁵

A theory that explains meanings as entities caused by objects is called a 'causal theory of perception' (CTP). So, to make the connection with the former section about computations; the objects that we perceive in any way through our senses cause mental representations, which are sentences or words in the LOT. The meaning of a mental representation is caused by the object it represents. A CTP should explain how this process works.

1.2.2 The Causal Theory of Perception

Basically the CTP says that mental representations are caused by the objects they represent. I will discuss some theories based on the work of Fodor. Note that 'represents' or a 'mental representation' are the same notions as 'meaning' or 'content'. The first causal theory I will discuss is what Fodor calls: the 'crude causal theory' (CCT).¹⁶ The CCT could be written as: "X" represents X iff X causes "X", and is actually the same as the basic notion of a CTP I stated above. However, as Crane points out, this theory is too naive, because a mental representation "X" is not necessarily caused by X.¹⁷ E.g. you can think of the sun without seeing it, or you can see the sun and think of vacation. So X does not necessarily cause "X" and "X" is not necessarily caused by X. What is wrong with the crude causal theory is that if meaning is based on the notion of 'causes', then there would be too much meaning in the world. For example, 'vacation' would mean the sun or anything else that causes 'vacation' thoughts.

What is missing in the CCT is a lawlike connection between the representation and the object. Crane distinguishes between two uses of 'represents' (or 'means').¹⁸ First, 'a red light

¹³ Imagine a twin-earth that is identical to our planet earth except for the fact that water has another molecular composition; it is not H₂O but XYZ. If you and your identical-self on twin-earth think of water (both without knowing anything about molecular compositions), then the content of your thoughts is the same. However, on earth this thought means H₂O and on twin-earth this thought means XYZ. So the environment determines the meaning of 'water'. - Putnam, H. *Representation and Reality*, pp. 30-3.

¹⁴ A little note on Rey's use of Putnam's thought experiment: Putnam argues against the ideas that there are semantic representations in the brain and that all our concepts are decomposable in such semantic representations. (Putnam, H. *Representation and Reality*, p. 5.) Rey uses Putnam's thought experiment only to show that there are at least *some* environmental aspects that are important to the agent in determining the meaning of a term.

¹⁵ Rey, G. *Contemporary Philosophy of Mind*. p. 241.

¹⁶ Fodor, J. A. *Psychosemantics*, ch. 4.

¹⁷ Crane, T. *The Mechanical Mind*, p. 176.

¹⁸ *Ibid.*, p. 175-7.

means stop' is based on a convention, because 'red' does not necessarily mean stop. Second, 'these spots mean measles' is based on the natural properties of the spots and measles and not on convention. Spots and measles have a lawlike connection or to put it differently: spots are a natural sign of measles. The CTP becomes: "X" represents X iff "X" reliably indicates X.¹⁹ Here 'reliably indicates' is just another phrase for 'natural sign', and means that there is a lawlike connection between "X" and X.

1.2.3 The representational component

At the beginning of the part about the representational component of CRTT I asked the following question: how can syntactically specified representations have specific meanings? In the light of what I have said, the answer is that the meaning of a representation is based on the connection that the representation bears with the object it represents. The 'connection' between object and representation is not just causal, but seems more like a lawlike connection. The answer I just gave is an external approach to a theory of meaning, so there may be other (internal) approaches to a theory of meaning within the CRTT. I used, following Rey, the twin earth thought experiment of Putnam to show that a completely internal approach is wrong. I neither defended the internal approach nor tried to argue against Putnam's argument, because I do not have the space for it in this thesis and more importantly, I focused on the external approach because I would like to discuss some problems with it in §2.

1.3 Final remarks

To summarize, together the computational and the representational component form the CRTT. The importance of the representational component is that it provides an explanation of how meaning gets into the mind.

So far I have ignored most problems with the CRTT because I only wanted to sketch what the theory tries to explain and how it tries to do so. Possible ways to attack CRTT on the computational component are to argue that its naturalistic, realistic or mentalistic assumptions are wrong or to argue that the idea of a LOT is wrong because there is no structure in the brain that consists of semantic representations.²⁰ In the next section I will look at problems that occur in the representational component of CRTT.

¹⁹ Ibid., p. 177.

²⁰ Putnam, H. Representation and Reality, p. 5.

2. Problems in the Causal Theory of Perception

The causal theory of perception is the part of the CRTT that explains how our thoughts acquire meaning. The definition “X” represents X iff “X” reliably indicates X (see §1.2.2), is not sufficient enough to explain meaning.²¹ This definition of CTP generates a problem. In the following I will describe that problem and a possible solution to it, and show, by giving examples, why this solution does not solve all the problems.

2.1 The disjunction problem

Remember Crane’s definition of the CTP: “X” represents X iff “X” reliably indicates X. There are two ways to understand the part *reliably indicates*. First, ‘reliably indicates’ could be seen as a *necessary condition* for representing.²² In this case “X” cannot represent X in the absence of X. This will have the consequence that we cannot mistake a goat for a sheep, because we cannot think “sheep” in the absence of a sheep. Crane calls this the misrepresentation problem.

Second, ‘reliably indicates’ could be seen as a *sufficient condition* for representing.²³ In this case “X” can represent X or Y or Z. If an object X, Y or Z causes the representation “X” then each of them reliably indicates “X”. For example, a “sheep” representation can reliably be caused by a sheep but also by a goat-in-certain-circumstances. The goat-in-certain-circumstances is a reliable indication so “sheep” means sheep or goat-in-certain-circumstances. Fodor’s example, where thoughts about Melbourne cause “platypus”, is in effect the same. The difference is that there are no actual objects involved in Fodor’s example. So if representations cause other representations, then the first representations reliably indicate the second representations. The problem then is that, for example, “sheep” could mean sheep, goat-in-certain-circumstances or anything that causes “sheep”. If “sheep” means anything, then the theory that is supposed to explain what “sheep” means does not work, because if everything can mean anything, then the concept of meaning dissolves. This is the disjunction problem.

Both the misrepresentation problem as well as the disjunction problem rule out error. The misrepresentation problem makes mistakes impossible, because error could not occur. The disjunction problem makes mistakes impossible, because cases of error are not cases of error according to the theory. So in both cases error could not occur according to the theory while it certainly occurs.

The Asymmetric Dependency Theory (ADT) seems to be a promising solution for this problem. In the following I will use ‘the disjunction problem’, ‘misrepresentation problem’, ‘problem of error’ and ‘misrepresentations’ as all referring to the problem I just described. Although these terms may not all exactly mean the same, these differences are not important to my point so I will use them interchangeably.

²¹ See:

- Crane, T. *The Mechanical Mind*, p. 178.

- Adams, F. “Fodor’s Asymmetrical Causal Dependency Theory of Meaning”, URL =<
<http://host.uniroma3.it/progetti/kant/field/asd.htm>>. Section I.

- Fodor, J. A. ‘Information and Representation’, pp. 100-1.

- Fodor, J. A. *Psychosemantics*, pp. 101-2.

²² Crane, T. *The Mechanical Mind*, p. 181.

²³ Ibid.

2.2 The Asymmetric Dependency Theory

The Asymmetrical Dependency Theory (ADT) is a solution, invented by Fodor, which should explain how we could establish a proper connection between a symbol and a representation without causing the disjunction problem. Adams²⁴ has reconstructed Fodor's version of ADT based on different articles by Fodor.²⁵

1. 'Xs cause "X"s' is a law,
2. For all Ys not = Xs, if Ys qua Ys actually cause "X"s, then Y's causing "X"s is asymmetrically dependent on X's causing "X"s,
3. There are some non-X-caused "X"s,
4. The dependence in (2) is *synchronic* (not diachronic).²⁶

Condition one is about natural signs. It says that "X" is a natural sign of X. The second condition is crucial, because it allows for non-Xs to be able to cause "X". This 'wrong' causation, in the sense that it does not fulfil condition (1), depends on a right causation, which does fulfil condition (1). I will use an example of Fodor to explain this. The representation "platypus" can be caused by an actual platypus (case a), but also by, for example, thinking about Melbourne (case b). Case a fulfils condition (1), but for case b we need condition (2), which says that it is only possible that thinking about Melbourne causes "platypus" if platypus causes "platypus". Furthermore, this last step does not work the other way around; case b is not dependent on case a. Therefore it is called the 'asymmetric dependency theory'.²⁷

Condition (3) is just meant to indicate that here are non-X-caused "X"s. Condition (4) means that if the law in condition (1) is broken, then all the cases that were depending on them are broken too. Furthermore, condition (4) is meant to prevent weird cases from happening, because if the dependence in (2) were diachronic, then this situation could occur: "Suppose that one comes to apply "X" to dogs, but only by means of observations of foxes. This would be a weird case of "learning", but if things were to go this way, one would not want "X" to mean fox."²⁸

So the causal theory of perception can now be defined as the ADT. It explains how normal cases of representing work and thereby it explains what misrepresentations are.

2.3 Problems with the Asymmetric Dependency Theory

ADT is a promising solution, however there are some problems that it cannot account for (yet). I will discuss three problems and then, in §4, elaborate on a more fundamental problem. The first two problems are that ADT has no conclusive explanation on (1) how names can indicate a person and (2) how there can be representations about non-existing objects. The third problem concerns lookalikes.

²⁴ Adams, F. "Fodor's Asymmetrical Causal Dependency Theory of Meaning", URL =<
<http://host.uniroma3.it/progetti/kant/field/asd.htm>>.

²⁵ According to Adams this could be found in Fodor 1987, 1990c and 1994.

²⁶ Ibid., Section II.

²⁷ Fodor, J. A. 'Information and Representation', p. 99.

²⁸ Adams, Fred and Aizawa, Ken, "Causal Theories of Mental Content", Section 3.4.

2.3.1 Names

According to the first condition of ADT there must be a lawlike connection between a representation and an object. But in the case of names there are no laws that apply uniquely to, for example, Aristotle in virtue of being Aristotle.²⁹ The problem then is: how can a name mean an individual? How can "Aristotle" mean Aristotle?³⁰ Fodor tries to solve this problem by arguing that the property of being Aristotle is lawfully connected with "Aristotle".

"[W]hat makes 'Aristotle' mean *Aristotle* in my mouth is a nomic relation between certain of Aristotle's properties; viz., between his property of *being Aristotle* and his property of *being (tenselessly) disposed to cause 'Aristotle' tokens in me.*"³¹

Fodor says that there is a nomic relation between two properties of Aristotle: (a) the property of being Aristotle and (b) the property of being disposed to cause 'Aristotle' tokens. But, as Adams points out, what if there are more individuals with the same 'Aristotle properties' (like clones)? Then they would all cause the same "Aristotle" token and consequently there would be no difference in referring to one rather than another Aristotle. However, this argument is, I think, not problematic for Fodor. It might be an undesirable consequence, but if Aristotle were to be cloned two times, then how could "Aristotle" mean the first cloned Aristotle rather than the second clone? It seems very intuitive to me that "Aristotle" means both of them. However, what if the clones develop differently over time? They could become different persons. Would "Aristotle" still mean both clones? Or has their property of being Aristotle changed? That leads to the question whether the property of being someone could change or not. I will not go into that question, since there is another more important point; I think that if two people are identical, then they necessarily *have* an identical property of 'being who they are'. If two people are non-identical then they necessarily *do not have* an identical property of 'being who they are'. And if there were two identical persons, then it would be, as I argued above, not strange, although not utterly unproblematic either, to refer to the same person with one name.

Another issue raised by Adams is that the name "Aristotle" seems to refer only to the property of being Aristotle, instead of the man or individual Aristotle. I think that Adams would be right to raise this issue if it were possible to consider the property of being Aristotle apart from the individual Aristotle. However, I think that having the property of being Aristotle is only possible if you are the individual Aristotle. So "Aristotle" refers to the individual Aristotle by means of the property of being Aristotle. Without that property, Aristotle would not be Aristotle anymore.

A more serious doubt that Adams raises is about the existence of properties like 'property of being X'. I think that this is a reasonable doubt, because how could we know whether such a property exists? It has explanatory strength but that does not prove its existence. Maybe the property could be understood as a theoretical property. However that would not tell us anything about how a name could mean an individual and how that connection is established in reality. So Fodor's solution to the problem of names is merely based on an assumption.

To conclude: Fodor does not provide a conclusive answer to the question of how a name can mean an individual. He assumes a special property, the 'property of being X', which should solve the problem, but we do not know if such a property really exists. And if it does, then it has

²⁹ Adams, F. "Fodor's Asymmetrical Causal Dependency Theory of Meaning", Section 4.

³⁰ Fodor, J. A. *The Elm and the Expert*, p. 110.

³¹ Fodor, J. A. *The Elm and the Expert*, pp. 118-9.

still some difficulties to overcome. However, Fodor can always say that his theory just does not (yet) explain how names can have meaning. The scope of his theory reduces, but it does not mean that his theory is completely refuted.

2.3.2 Non-existing objects

According to ADT every representation needs to be instantiated. However, how could that be for non-existing objects? We can think about unicorns, Santa Claus, Zeus, etc. while they are all non-existing objects, which cannot have instantiated the causal relation that the ADT requires in its first clause. A reply of Fodor to this problem is saying that unicorns etc. are non-primitive.³² A non-primitive would be a description that can cause a primitive representation. Fodor is not quite clear on how this works, but I think that the non-primitive description consists of primitive parts, e.g. 'horn' and 'horse', that can cause the right representations to instantiate the primitive concept, e.g. "unicorn".

Baker argues that there could be a twin-earth with an identical you, only on twin-earth there are real unicorns. So it follows that you (on earth) and your doppelgänger have the same syntactic symbol "unicorn" but yours is semantically non-primitive while the symbol of your doppelgängers is semantically primitive.³³ According to Baker this severs the connection between the syntax and semantics of the symbol, because the same symbol cannot have different semantic properties.³⁴ This implicates that Baker is an externalist, because she assumes that the meaning of a symbol also depends on its semantic properties. And if a symbol could have more than one semantic property, then there is no way of knowing whether that symbol means one or the other semantic property.

Fodor replies by saying that "the *point* of introducing concepts by definition is that it allows for syntactically primitive expressions that are semantically equivalent to descriptions."³⁵ Fodor is an internalist, which means that he thinks that meaning is reducible to syntax and does not depend on semantic properties of a symbol. Therefore, he can say that a symbol can have more than one semantic property, because the meaning of a symbol depends only on the syntax of the symbol. So you and your doppelgänger do have the same syntactic symbols with different semantic properties. The description (non-primitive) is equivalent to the semantic property.

Thus the difference between Baker and Fodor in understanding non-existing objects is that Baker treats them from an externalistic and Fodor from an internalistic point of view. Fodor's solution to this problem depends on the truth of the internalist view, because if the meaning of a symbol only depends on its syntax, then it can have different semantic properties.

I will not pursue the internalist/externalist debate on this matter. The important point is that this problem does not just depend on the ADT and that other assumptions play a role too.

2.3.3 Lookalikes

Earlier, when I discussed the external approach, I followed Rey in using Putnam's twin-earth thought experiment to show that meaning is not just in the head. I also said that this was an interpretation of Putnam that Putnam would not agree on. The twin-earth thought experiment could be used as a counter example of the ADT, which is more in the line of Putnam's intentions.

³² Loewer, B. and Georges Rey, *Meaning in Mind*, p. 259.

³³ 'Semantically primitive' means that the concept can be reduced to a single symbol in the LOT and 'semantically non-primitive' means that the concept is composed of different semantic primitives.

³⁴ Baker, L. R. 'Has Content Been Naturalised?', pp. 21-2.

³⁵ Loewer, B. and Georges Rey, *Meaning in Mind*, p. 259.

The problem for the ADT is that it cannot explain the difference in meaning of “water” for Earthlings as opposed to Twin-Earthlings.³⁶ Dennett puts this problem in a more general formulation: “How does Fodor establish that, in his mental idiolect, “horse” means *horse*—and not *horse-or-other-quadruped-resembling-a-horse* (or something like that)?”³⁷ Or in Adams words: what keeps “X” from meaning X or X-lookalike?³⁸

Baker³⁹ gives an example of this imputation by coming up with a case of a misrepresentation without an asymmetrical dependency. Imagine a robot-cat that looks exactly like a real cat. Suppose that person S learns the mental representation “cat” by looking at robot-cats. After a while S sees the real cat for the first time, which causes the mental representation “cat”. The question then is: how should this cat-caused token be interpreted?

“There seem to be three possibilities for interpreting the cat-caused token. Either (i) it is a correct representation of a cat (and the robot-caused tokens are misrepresentations of cats); or (ii) it is a misrepresentation of a robot-cat (and the robot-caused tokens are correct representations of robot-cats); or (iii) it is a correct representation of the disjunctive property cat-or-robot-cat.”⁴⁰

According to the first interpretation the cat-caused token means ‘cat’. The robot-cat tokens are asymmetrically dependent on the single cat token. However that is not right because robot-cats were first tokened and it would be impossible to be asymmetrically dependent on un-tokened instances. The second interpretation says that the cat-caused token means robot-cat. But this is also wrong, because it ignores relevant counterfactuals. Real cats cause the same tokens as robot-cats so there is no asymmetry between real cats and robot-cats caused tokens. It is only an accident that the robot-cats were encountered first. The third interpretation says that it means both cat and robot-cat, which is an occurrence of the disjunction problem that ADT tries to solve.

Baker leaves these three undesirable options for Fodor and in her paper ‘On a Causal Theory of Content’ she concludes this argument, by stating that the ADT does not solve the problem of misrepresentations.⁴¹

Fodor chooses option three and recalls Dretske’s variation of the twin-earth thought experiment to vindicate his decision. Suppose that H₂O and XYZ are both on twin-earth and someone learns “water” only from samples of H₂O, then Fodor’s and Dretske’s intuitions are that “water” means H₂O or XYZ. This example is similar to a world with cats and robot-cats so option three is valid.⁴² But Fodor also says that it is still correct to say that, on Putnam’s version of twin-earth with only XYZ, “water” is true of XYZ and not of H₂O.⁴³

How does defend Fodor these intuitions? First, let me explain the difference between narrow and broad content. “A *narrow* content of a particular belief is a content of that belief that is completely determined by the individual’s intrinsic properties. An *intrinsic* property of an individual is a property that does not depend at all on the individual’s environment.”⁴⁴ Broad content is content that is determined by the environment, for example in the twin-earth thought experiment where the meaning of “water” depends on the presence of XYZ or H₂O. Furthermore, the narrow content determines, based on the environment, the broad content. So I can have

³⁶ Adams, F. “Fodor’s Asymmetrical Causal Dependency Theory of Meaning”, Section 4.

³⁷ Dennett, D. C. *The Intentional Stance*, p. 309.

³⁸ Adams, F. “Fodor’s Asymmetrical Causal Dependency Theory of Meaning”, Section 4.

³⁹ Baker, L. R. ‘On a Causal Theory of Content’, p. 167.

⁴⁰ *Ibid.*, p. 168.

⁴¹ *Ibid.*, p. 169.

⁴² Fodor, J. A. *A Theory of Content and other Essays*, p. 104.

⁴³ *Ibid.*

⁴⁴ Brown, Curtis, “Narrow Mental Content”, *The Stanford Encyclopedia of Philosophy*

narrow content about “water” that, when I am in an environment where water is, determines the broad content, or the meaning, of “water”: H₂O. Important in this process is the environment where the agent first acquired his beliefs and mental states. Twin-me still thinks of XYZ if he would see water on earth, because he first acquired the beliefs about water on twin-earth where “water” means XYZ.

By using this distinction between narrow and broad content Fodor can say that “cat” has a narrow content that based on the environment, in this case the presence of cats or robot-cats, determines the broad content which could mean cats or robot-cats. This works because S has acquired both beliefs in the environment of cats and robot-cats.

Apart from the question whether the distinction between, and the use of, narrow and broad content is valid, the problem with this solution is, I think, that it doesn’t say anything about error. If S learns what “cat” means by looking at robot-cats and later discovers that some robot-cats are real cats, then there must be a way that she can say that she was mistaken. What Fodor does is saying that she was not mistaken, and he shows us how “robot-cats” can mean both robot-cats and cats. So the same happens as with the disjunction problem: error disappears.

2.4 Final remarks

I have discussed three problems for the ADT that Fodor tries to solve. I have argued that in the case with names, Fodor does not give a conclusive answer and the answer that he provides depends basically on the existence of a property that he assumes. If that property exists then his argument gains strength, however there remain some problems. In the ‘non-existing objects case’ Fodor’s defence was based on the existence of non-primitive semantic properties. However that included the assumption that meaning is only based on syntax, which is an internalistic point of view. So whether Fodor is right depends on his internalistic standpoint. I did leave the discussion there, but I think it shows that ADT is not sufficient on its own and needs other assumptions or theories in order to work properly. At last I considered lookalikes, which could cause a misrepresentation without asymmetric dependency. Fodor solved this problem with the distinction between narrow and broad content, however his solution did not recognise error and added to that, the theory about narrow and broad content is also an extra assumption to make the ADT work properly.

All three cases do not refute the ADT, however they show that the theory needs a lot of extra controversial assumptions to account for problematic cases. I did not show that the ADT is false, but I did not show that it is unproblematic either. The more additional theories Fodor uses, the more strength ADT loses, and what is the usefulness of a theory that only works in standard cases? Despite that, Fodor can always say that ADT does not account for these problems yet. But as long as he does not have a clear answer to these problems, ADT should not be taken too seriously.

3. The deeper problem

The ADT explains that misrepresentations are dependent on proper representations. However it does not explain how we can distinguish erroneous from non-erroneous misrepresentations and thereby it does not explain how misrepresentations emerge. Both Fodor and Crane recognise that the theory falls short on those issues. Crane: "I do not see how this [the ADT] gives us a deeper understanding of how mental representation actually works."⁴⁵ Fodor: "we need a story about error not because error raises the disjunction problem, but just because we need a story about error. And it would be very desirable if the story we tell explained why calling something a mistake has the normative force that it does."⁴⁶

In the following I will scrutinise these issues and explain what a fundamental problem for a CTP is. Then I will discuss a probable solution to this problem.

3.1 Normativity in perception

To be able to distinguish erroneous from non-erroneous misrepresentations and explain how they both emerge, we could try to give an exact description of error in perception. However a naturalistic description of the process of perception and describing where and how errors could occur would not suffice. If I explain that error emerges, for example, 'when light that reflects on a cow is somehow altered before it hits the cells in my retina so that I don't see a cow, but a horse,' which is a misrepresentation, then I have explained the cause of the error. But that explanation already assumes error. If I change the semantic phrases, which are underlined, with non-semantic and non-intentional phrases, then the same explanation would go like this: 'when light that reflects on object X is somehow altered before it hits the cells in my retina and causes me to have a representation, then that object X causes a representation "X".' But this explanation doesn't tell me whether "X" representing X is a misrepresentation or not. I see an object and I have a representation, but how am I to know that the representation is correct? Thus, either the explanation contains semantic assumptions and falls short on circularity, or the explanation does not tell us anything about whether the representation is correct or false.

But why is it necessary to describe a theory of perception in non-semantic and non-intentional terms? This is because if we would use semantic and intentional terms in explaining how representations can have semantics, then we already assume what we want to explain. And that is circular.

Another way of expressing the problem is by saying that representations have a normative property. This property says in which cases the representation is true or false. A representation is true if it represents the world as it is, and it is false if it does not. So the representation "cat" would be true if there is a cat. The question then is, what is that normative property? And again, this should be expressed in non-semantic and non-intentional terms.

3.2 Solution?

One possibility would be accepting circularity, but I think that would be highly undesirable because it leaves the theory with a big problem. The other option would be to try and find a way to explain in non-semantic terms what misrepresentations are.

⁴⁵ Crane, T. *The Mechanical Mind*, p. 185.

⁴⁶ Fodor, J. A. 'Information and Representation', p. 102.

To begin with it should be clear that we could not use semantics in explaining error, because that would lead to circularity. However, without semantics there is no error, because error occurs only when there is semantics. Two symbols do or do not mean the same until you know what the symbols mean. So there simply is no error until there is semantics involved.

The only way to establish a theory of error would be to find a substitute for semantics so that we can say without semantics whether the connection between representation and the object is erroneous or not. In 'Information and Representation' Fodor provides a rough outset of what such a theory may look like.⁴⁷

3.2.1 Abortive actions

According to Fodor the ADT explains what it is for a symbol to express a property, but it does not explain how a symbol can be applied to an object.⁴⁸ So he is saying that the ADT is good but that there also needs to be a theory that explains how the application of symbols to objects work. Fodor says that if you apply a symbol to an object then you put a token of this symbol and the connection to the object in your belief-true box, e.g. the beliefs that "Bossy is a platypus" and "platypai lay eggs". Based on the various items in the belief-true box you can act. If you get the desire to eat a platypus egg omelette, then you will go looking in the garden where Bossy lives. However you may have been wrong in believing that Bossy is a platypus, it was a misrepresentation, so you do not find platypus eggs in your garden.⁴⁹ Fodor summarises this as follows: "According to this view, the naturalistic answer to "what's wrong with false beliefs?" is primarily that false beliefs lead to abortive actions."⁵⁰ So, a representation is false if the beliefs based on it lead to abortive actions, which are actions that do not lead to the intended goal of that action.

Besides the question if this is an undesirable form of functionalism,⁵¹ this seems to be a very vague way of determining whether a representation is a misrepresentation or not, because maybe your actions are such that they failed for some other reason. E.g. Bossy could actually be a platypus that just does not lay eggs or you did not look well enough for eggs. So an abortive action does not automatically indicate a misrepresentation. Thereby, a misrepresentation does not automatically lead to an abortive action either. E.g. the representation of Bossy was wrong; in fact she is a cow, but because the platypus of your neighbours escaped to your garden and laid eggs, you still find them and have your omelette. So the misrepresentation did not lead, in this case, to an abortive action, which off course could occur at a later point.

Another problem with Fodor's proposal is that, according to Crane, the desire is expressed in semantic terms, which leads to circularity.⁵² Crane expresses Fodor's proposal⁵³ as follows: "A belief B represents condition C if and only if actions which would be caused by B and a desire D would bring about the state of affairs that represented by D were C to obtain."⁵⁴ Belief B is caused by objects and condition C is about how the belief says the world is. But if we want to know whether belief B represents condition C, or in other words, if we want to know the semantic properties of B and C, then we may not assume semantics in the explanation after the

⁴⁷ He restricts himself to misrepresentations that are 'mislabelings', e.g. when a horse causes "cow". The outset is nota about non-existing objects or representations caused by other representations.

⁴⁸ Fodor, J. A. 'Information and Representation', p. 102.

⁴⁹ Assuming that there are no other eggs of other platypai in your garden.

⁵⁰ Fodor, J. A. 'Information and Representation', p. 102.

⁵¹ I don't have the space to explain why Fodor doesn't want his theory to be functionalistic and why he still considers this as a valid solution. See also footnote 8 of 'Information and Representation' in Lycan and Prinz, p. 104.

⁵² Crane, T. *The Mechanical Mind*, p. 189.

⁵³ Crane doesn't say that this expresses Fodor's proposal, but I think that it is a proper description of it.

⁵⁴ Crane, T. *The Mechanical Mind*, p. 189.

phrase 'if and only if', because that would be circular. But in fact there are semantic notions involved in defining the 'desire D that would bring about the state of affairs represented by D', so this solution is circular too.

Again it should be noted that the idea of abortive actions is only a rough outset given by Fodor. But despite that, I think that the argument of Crane that the idea of abortive actions leads to circularity is hard to overcome, because in any way that you want to compare your desires with the result of your actions, you must appeal to semantics.

3.3 Final remarks

Both the ADT and the 'abortive actions' of Fodor do not give a satisfactory explanation of normativity in perception. They both lead to circularity, which is an unacceptable outcome. Furthermore, the ADT has a hard time with dealing with difficult cases. Whether they are or are not conclusive in rejecting the ADT does not matter a lot when there is no explanation of normativity in perception.

4. Conclusion

I will first summarise the highlights and then I will answer the main question. I started by explaining the CRTT (§1). Then I mentioned two approaches to explain how our thoughts could have meaning, an internal and an external (§1.2.1). The internal would not work as Rey showed with Putnam's twin-earth experiment. I followed the external approach, which led to the disjunction problem and the problem of error (§2.1). Fodor proposes the Asymmetric Dependency Theory (§2.2), which says, to put it bluntly, that error is parasitic on truth. And despite the fact that this theory has a hard time dealing with problematic cases, it does not explain error at all, because it falls short on circularity (§3.1). Then I argued that the only possible way to overcome circularity is to find a substitute for semantics by which we can tell whether a representation is erroneous or not. I discussed Fodor's 'abortive actions' theory, which turned out to be also circular (§3.2.1).

When we look at the theories and solutions I have scrutinised, then I think it is fair to say that the disjunction problem is not solved and that the options and solutions I discussed are all problematic or impossible. First of all, ADT does not explain how error occurs. It merely describes it as 'parasitic on truth'. Thereby, ADT has a hard time dealing with various kinds of representations and needs extra theories or assumptions to account for them. These extra theories are also not uncontroversial. Therefore, I think that ADT is, for now, an inadequate theory when it comes to handling the disjunction problem. However it is not refuted either. So it is possible that it could work with some major fixes. Hence my conclusion is that, based on the theories I have discussed, Fodor has failed to solve the disjunction problem and the problem of normativity in the Causal Theory of Perception.

Literature

Books

Crane, T. *The Mechanical Mind: a philosophical introduction to minds, machines and mental representations*. Second edition. Oxon : Routledge, 2003. (First edition : 1995)

Dennett, D. C. *The intentional stance*. Seventh edition. Cambridge, Mass.: The MIT Press, 1998. (First edition : 1989)

Fodor, J. A. *A Theory of Content: and other essays*. Cambridge, Mass.: Bradford Books, 1990.

Fodor, J. A. *Psychosemantics: the problem of meaning in the philosophy of mind*. Cambridge, Mass.: Bradford Books/The MIT Press, 1987.

Fodor, J. A. *Representation and Reality*. Cambridge, Mass.: Bradford Books/The MIT Press, 1988.

Loewer, B. and Georges Rey (ed.). *Meaning in Mind: Fodor and his critics*. Cambridge, Mass.: Blackwell, 1991.

Lycan, William G. and Jesse J. Prinz (ed.). *Mind and Cognition: an anthology*. Third edition. Oxford : Blackwell Publishing Ltd., 2008.

Rey, G. *Contemporary Philosophy of Mind*. Cambridge, Mass.: Blackwell, 1997.

Articles

Baker, L. R. 'Has Content Been Naturalised?,' in Barry Loewer and Georges Rey (ed.), *Meaning in Mind* (Cambridge, Mass.: Blackwell, 1991), pp. 17-32.

Baker, L. R. 'On a Causal Theory of Content,' in *Philosophical Perspectives, Vol. 3, Philosophy of Mind and Action Theory* (1989), pp.165-186

Fodor, J. A. 'Information and Representation,' in William G. Lycan et al. (ed.), *Mind and Cognition: an Anthology* (Oxford : Blackwell Publishing Ltd., 2008), pp. 95-104.

Fodor, J. A. 'Observation Reconsidered,' in *Philosophy of Science*, Vol. 51, No. 1. (Mar., 1984), pp. 23-43.

Fodor, J. A. 'Why There Has to Be and How There Could Be a Private Language' in William G. Lycan et al. (ed.), *Mind and Cognition: an Anthology* (Oxford : Blackwell Publishing Ltd., 2008), pp. 155-70.

Quine, W. V. 'Quantifiers and Propositional Attitudes' in *The Journal of Philosophy* Vol. 53, No. 5 (Mar. 1, 1956), pp. 177-187.

Websites

Adams, Fred, "Fodor's Asymmetrical Causal Dependency Theory of Meaning", URL =<<http://host.uniroma3.it/progetti/kant/field/asd.htm>>. (last visited 26 August 2012)

Adams, Fred and Aizawa, Ken, "Causal Theories of Mental Content", *The Stanford Encyclopedia of Philosophy (Spring 2010 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2010/entries/content-causal/>>.

Aydede, Murat, "The Language of Thought Hypothesis", *The Stanford Encyclopedia of Philosophy (Fall 2010 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2010/entries/language-thought/>>.

Brown, Curtis, "Narrow Mental Content", *The Stanford Encyclopedia of Philosophy (Fall 2011 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2011/entries/content-narrow/>>.

Jacob, Pierre, "Intentionality", *The Stanford Encyclopedia of Philosophy (Fall 2010 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2010/entries/intentionality/>>.

McKay, Thomas and Nelson, Michael, "Propositional Attitude Reports", *The Stanford Encyclopedia of Philosophy (Winter 2010 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/win2010/entries/prop-attitude-reports/>>.

Neander, Karen, "Teleological Theories of Mental Content", *The Stanford Encyclopedia of Philosophy (Spring 2012 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2012/entries/content-teleological/>>.

Horst, Steven, "The Computational Theory of Mind", *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2011/entries/computational-mind/>>.