

## *The psychology of meaning*

# Lecture 1: Psychological explanations

### 1. Introduction

- The aim of this course is to explore the Gricean view on interpretation, in general, and develop a detailed account of *quantity implicatures*, in particular.
- Examples of quantity implicatures:
  - (1) a. Many of the nurses were drunk. *“scalar implicature”*  
    ↪ Not all the nurses were drunk.
  - b. If you scratch my back, I’ll scratch yours. *“conditional perfection”*  
    ↪ If you don’t scratch my back, I won’t scratch yours.
  - c. You may have an apple or a pear. *“free choice permission”*  
    ↪ You may have an apple.  
    ↪ You may have a pear.
- The Gricean view on interpretation is intimately connected to what Grice calls “philosophical psychology”. In this lecture we will ask ourselves what that might be.
- However, we will try to answer this question by focusing other philosophers’ views, especially Daniel Dennett’s (1987a, 1987b), because these tend to be more explicit (and therefore more vulnerable) in various respects.

### 2. Levels of analysis of information processing systems

Suppose it can be shown that addition is easier than multiplication. Then how do you explain that a given computer needs more time for doing multiplication than for doing addition? *Not* at the hardware level, obviously. The same holds, *mutatis mutandis*, for other kinds of behaviour.

“[Trying] to understand perception by studying only neurons is like trying to understand bird flight by studying only feathers: it just cannot be done.” (Marr 1982: 27)

Marr (1982) famously distinguishes between:

- The computational level: What is computed? (e.g. multiplication)
- The representational/algorithmic level: How is the problem represented and the computation carried out? (e.g. long multiplication or peasant multiplication)
- The implementational level: What kind of hardware executes the algorithm?

“These three levels are coupled, but only loosely.” (p. 25)

Dennett makes a somewhat similar distinction between:

- The intentional stance (beliefs and desires): this is at about the same level as Marr’s computational level, but with important restrictions.
- The design stance (~ Marr’s algorithmic level)
- The physical stance (~ Marr’s implementational level)

For our purposes (and for Dennett’s, as well) the intentional stance is the most important one.

- Dennett emphasises that there may be a certain amount of slack between levels: When viewed at a deeper level, a system may fall short of the requirements specified at a higher level.
- Therefore, an intentional analysis may be *idealised* in certain respects. This can cause methodological dilemmas: How to distinguish between an analysis that is wrong and one that is merely idealised? Cf. Chomsky’s (1965) distinction between *competence* and *performance*.
- Folk psychology adopts the intentional stance (though it allows admixtures from the other stances, too): We treat each other as intentional systems.  
Hence, the intentional stance is used in everyday life as well as in the behavioural sciences.

### 3. Ontological concerns

What is the ontological status of concepts like “belief”, “desire”, and so on? It is rather obvious that beliefs don’t exist in the same way as chairs or apples do. But do they exist at all?

Dennett tries to position himself very carefully between ontological *realism* and *interpretationism* (or *instrumentalism*), but Grice’s ontological views are more liberal:

My taste is for keeping open house for all sorts and conditions of entities, just so long as when they come in they help with the housework. [...] To fangle a new ontological Marxism, *they work therefore they exist*, even though only some, perhaps those who come on the recommendation of some form of transcendental argument, may qualify for the specially favoured status of *entia realissima*. (Grice 1975/1991: 131)

#### 4. Rationality

According to Dennett, Grice, Davidson, and many other philosophers, adopting the intentional stance entails that the behaviour to be explained is rational.

“[The] use of language is one among a range of forms of rational activity.” (Grice 1989: 341)

This assumption raises various issues:

- How plausible is it to assume by default that people (say) are rational? There are quite a lot of experimental data which have been argued to show the opposite (e.g., Sutherland 1992, Marcus 2008).
- What is rationality, in the first place? Aren't the standard canons of rationality way too strong? If yes, what is going to replace them?

##### 4.1. How rational are we?

- Human reasoning seems to be quite fallible (e.g., syllogistic reasoning, conditionals).
- Reasoning about probabilities is hard for humans.  
E.g., the conjunction fallacy (Tversky and Kahneman 1983):

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Which is more probable?

- Linda is a bank teller.
- Linda is a bank teller and is active in the feminist movement.

- Framing (Tversky and Kahneman 1981):  
Condition 1:

Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

- If Program A is adopted, 200 people will be saved.
- If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.

Which of the two programs would you favor?

Condition 2:

- If Program C is adopted, 400 people will die.
- If Program D is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die.

#### 4.2. *Minimal rationality*

The standard view on rationality:

- Rationality is a necessary condition for agenthood.
- In principle, an agent *A* should undertake all and only actions that, given *A*'s beliefs, will help *A* to attain his goals.
- In principle, *A*'s beliefs are closed under deduction.
- In principle, *A*'s belief set is consistent.

Cherniak (1986) argues rather convincingly that requirements like these are *wildly* unrealistic:

- Various meta-logical theorems show that these ideals cannot be attained, even in principle. (E.g., Gödel's incompleteness theorem, Church's undecidability theorem.)
- Even if a proof system is decidable, it may be unrealistic for any *finite* agent. (E.g., establishing consistency for propositional theories.)
- There are considerable practical limits on what real agents can achieve with the cognitive equipment they *actually* have. (E.g., limited capacity of working memory, limitations imposed by the structure of long-term memory.)

Cherniak proposes to deal with these problems by weakening the classical rationality conditions along the following lines:

- An agent *A* should undertake *at least some* of the actions that, given *A*'s beliefs, will help *A* to attain his goals.

This is extremely weak, but the theory is strengthened by “background theories” of, e.g., feasible inferences and the structure of memory.

Worries:

- If Cherniak is right, there may be very little that rational agents must have in common. There seems to be little left of the classical concept of rationality.
- Concerning Dennett’s intentional stance: What Cherniak proposes, in effect, is that the notion of rationality is heavily constrained by an agent’s *design*. This raises the question of whether it is possible to neatly separate the intentional stance from the design stance.
- Come to think of it: the design level is probably constrained by the implementational level (e.g., memory). But this would imply that what is rational for a given organism is partly determined by its hardware.
- More briefly: Is it really possible to neatly separate the stances from each other?

## References

- Cherniak, C. (1986). *Minimal rationality*. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, Massachusetts: MIT Press.
- Dennett, D. C. (1987a). *The intentional stance*. Cambridge, Massachusetts: MIT Press.
- Dennett, D. C. (1987b). True believers: the intentional strategy and why it works. In: Dennett (1987a:13-42).
- Grice, H. P. (1989). *Studies in the way of words*. Cambridge, Massachusetts: Harvard University Press.
- Grice, P. (1975). Method in philosophical psychology (from the banal to the bizarre). *Proceedings and addresses of the American Philosophical Association*, (48): 23–53. Reprinted in and cited from Grice (1991:121-161).
- Grice, P. (1991). *The conception of value*. Oxford University Press.
- Marcus, G. (2008). *Kluge: the haphazard evolution of the human mind*. New York: Houghton Mifflin.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: W.H. Freeman and company.
- Sutherland, S. (1992). *Irrationality*. London: Constable & Co.
- Tversky, A. and D. Kahneman (1981). The framing of decisions and the psychology of choice. *Science* 211: 453–458.
- Tversky, A. and D. Kahneman (1983). Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychological review* 90: 293–315.