ISSUES IN PRAGMATICS (PLIN 3001) 2006-07

LEXICAL PRAGMATICS

9. Innateness revisited

1. Introduction

We started this term with Fodor's critique of the classical view of word meaning, and considered his argument that word meanings are in general simple, unanalysable concepts which do not decompose into clusters of simpler concepts. We then looked at what he saw as one of the main consequences of his account: that simple, unanalysable concepts are **triggered**, rather than learned by standard methods of hypothesis formation and confirmation, and are therefore in some sense innate. In Lecture 5, we considered a stronger and a weaker version of the **innateness-and-triggering** account:

(1a) The **concept itself** (e.g. RED) is innate and merely **activated** by experience (e.g. of red things).
(1b) We have innately-determined **concept construction mechanisms** which are triggered by experience and construct concepts that go well beyond the available evidence.

On the stronger version, which Fodor seems to have held until 1998, we are born with a set of innate **concepts**, or **ideas**, which are merely activated or released by experience. On the weaker version, we are born with a set of **concept-construction mechanisms** which create different **kinds** of concept when confronted with objects, events and properties from different **cognitive domains.** I ended Lecture 5 with three quotations from Fodor's 1998 book *Concepts* which suggest that he would now endorse this slightly weaker view:

'Since the sensorium isn't an idea, it is a fortiori not an *innate* idea. So, strictly speaking, the innate sensorium model for the acquisition of RED doesn't require that it, or any other concept, be innate'. (Fodor 1998: 142).

'The kind of nativism ... that an informational atomist has to put up with is perhaps not one of *concepts* but of *mechanisms.* (Fodor 1998: 142).

'Maybe there aren't any innate *ideas* after all'. (Fodor 1998: 143).

In the last three lectures, we've seen further evidence for the weaker version of nativism. We considered an account of lexical pragmatics which assumes that every hearer has an ability to construct **ad hoc concepts**, on the model of existing concepts, during the on-line

comprehension of almost any word. This account seems to fit much better with the weaker view that there are innate **mechanisms** rather than the stronger view that there are innate **concept**s: innate mechanisms would allow us to construct an indefinite number of new concepts, using existing formats, to deal with new experiences, rather than assuming that they are all present from birth. Today, I'd like to look a little more closely at this idea that there are different **kinds** of concept, with different, innately-determined **formats**, and show how it offers a middle way between radical-nativist and radical-empiricist views.

The idea that there are different kinds of concepts has surfaced in many different forms in both philosophy and psychology. Recently it's become increasingly popular in work inspired by evolutionary psychology, which claims that different kinds of concepts are found in different cognitive domains (see e.g. Leda Cosmides & John Tooby, 'Evolutionary psychology: A primer', available from Leda Cosmides' website). Different cognitive mechanisms are seen as having evolved at different times, in response to specific problems in the organism's environment: thus, the concepts and mechanisms found in one domain may have little in common with those in other domains. Typically, evolutionary approaches rest on the following types of argument and evidence (Dehaene 1997):

(2a) Possessing a certain ability would confer an **evolutionary** advantage.
(2b) There are **precursors** of this ability in other animal species.
(2c) It emerges **spontaneously** in children, independently of other abilities such as language.
(2d) It is not acquired by **slow, domain-general** mechanisms of learning.
(2e) It should have a distinct neural substrate.

This approach seems to offer a middle path between two extreme views of concept acquisition: the **empiricist** view that concept acquisition is largely environment-led, with concepts merely **copied** from the environment, and Fodor's early view that it is largely organism-led, with the environment serving merely to **activate** innately determined concepts. The hypothesis is that the mind is a sort of **adaptive toolkit** with a wide variety of mental modules which have evolved in response to particular environmental problems, and suggests that what is innately determined might be the formats for constructing (or learning) different **kinds** of concepts, with individual details being filled in from the environment. This might allow us to do justice both to Fodor's innateness arguments and to the intuition that he sometimes goes too far.

2. Domains as innately-determined theories or modules

At the end of Lecture 5, we gave a brief sketch of some of the evidence that the child is born with innate assumptions, or biases, which structure its development in different cognitive domains. In this section, I'll briefly summarise some of the main findings in this area that most developmental psychologists would currently agree on. They show a systematic abandonment of empiricist assumptions in the face of counter-evidence, and adoption of some version of a nativist account.

Naive physics

Empiricist approaches tend to assume that the child is born with no knowledge of the nature of physical objects, and no expectations about how such physical objects might behave (e.g. that they persist through time, have causal interactions with each other, continue to exist when we are not perceiving them, etc.). For the empiricist, the world is no more than a set of sense experiences (**sense data**), which depend on the perceiver. Piaget, one of the pioneers of developmental psychology, took this view:

'the universe of the young baby is a world without objects, consisting only of shifting and unsubstantial 'tableaux' which appear and then are totally reabsorbed, either without returning, or reappearing in a modified or analogous form' (Piaget, 1969:14).

Thus, Piaget did experiments to show that a child will not reach for a toy hidden under a blanket until the age of around 9 months. More generally, it will have no particular **expectations** about the behaviour of physical objects: nothing will **surprise** it.

More recent experimental work seems to show convincingly that children are born with certain biases or expectations about the behaviour of physical objects. For example, Spelke 1990 and Premack 1994 summarise a range of experiments which show that children from around 4 or 5 months are surprised when one solid object appears to pass through another. Children from around 6 months seem to interpret interactions among solid objects in causal terms: for example, when two billiard balls collide, they see one as the 'pusher' and one as the 'pushee', and expect the 'pushee' to move immediately and the speed of the pusher to affect the speed of the pushee. Mehler & Dupoux 1990 report experiments involving a screen with a window in the middle of it, which showed that children at 14 weeks expect a tall rabbit that disappears behind the screen to be visible through the window as it runs past, while a small rabbit is not expected to appear in the screen. These expectations are so strong that some

infants become ‘intensely agitated’ if their predictions are not met. Summarising a range of evidence, Baillargeon 1985 draws the following conclusion:

‘The general picture suggested by the present research is, thus, one in which the physical world of infants appears very similar to that of adults’ (Baillargeon, 1985:609).

Self-propelled objects

Among solid objects, there's an important distinction between those that are **self-propelled** and those that are not. A series of papers (by e.g. Premack 1990, Spelke & Carey 1996 and Leslie 1994) investigates how the child tells them apart: essentially, arbitrary changes of speed or direction suggest that an object is self-propelled. Self-propelled objects are then candidates for having **minds**, and providing input to the famous ‘**theory of mind’** module, or **mindreading** module.

Mindreading module

In the area of mindreading, too, recent experimental evidence seems to undermine the classical empiricist view. Piaget did experiments which seemed to show that the child remained ‘egocentric’ and unable to see the world from the perspective of another until around 5 or 6:

‘it is only after long training that child reaches the point … where he speaks no longer for himself but from the point of view of the other’ (Piaget: 1966:122).

One piece of evidence for this was the Swiss Mountain test, where the child was shown a model of three mountains with various distinguishing features: one with snow on the top, another with a house, and a third with a church. The child is allowed to examine the model from all angles, then a doll is placed at a certain position in the model, and the child is asked what the doll can see. According to Piaget, it was not till around the age of 6 that children could perform accurately on this task: until then, they simply reported what they themselves could see.

More recent experimental work provides evidence for a **mindreading module** which is itself made up of a cascade of sub-modules, maturing at different stages in the development of the child. One of the earliest is the **Eye Direction Detector** (Baron-Cohen 1995), which infers what someone is seeing or watching from the direction of their gaze. This starts developing from around the age of 6-8 months, contrary to what Piaget predicts on the basis of the Swiss Mountain test. Another deals with **goal-directed actions**. As Leslie points out, naive physics is

purely local ('contact mechanics'), whereas goal-directed actions involve relations between agents and states of affairs which are distant from them in space and time. Since these states of affairs may not actually exist (e.g. John may hide in the doorway because he **thinks** it's raining, even though it's not), and since non-existent events can't be causes, it's best to see the agent's actions as caused by his **mental representation** of states of affairs. Leslie sees these mental representations as proposition-like, e.g. like sentences in Fodor's language of thought.

Leslie is best known for his work on the sub-component of the **mind-reading module** which gives humans the ability to attribute **propositional attitudes** – intentions, beliefs, desires – to others in order to predict and explain their behaviour. This begins to develop during the second year of life. The earliest sign is the ability to engage in some form of inferential communication, which of course is interpreted in terms of intentions, and happens from 12 months on. Another early sign is the ability to **pretend** and to understand pretence, which emerges between 18 and 24 months (and which according to Leslie involves a metarepresentational ability). This sub-component of the theory of mind enables the child to interpret the behaviour of others as caused by beliefs, desires, intentions and other mental states. This sensitivity to the intentions of others plays an important role in vocabulary acquisition, as we'll see next week.

A much later development (around 3.5) is the ability to attribute **false beliefs** to others, as in the famous Sally-Anne task. Here, puppet A (Sally) hides a sweet in a box, then leaves the room; while A is out, puppet B (Anne) helps the child move the sweet to another box. When A returns, the experimenter asks the child, 'Where will A look for the sweet?' Up to age 3.5, children tend to assume that A will look in the new box, where the sweet actually is (despite the fact that A didn't know about the move); after 3.5, they realise that she will look in the old box, where the sweet was when she left the room. In other words, by the end of the fourth year, children can cope with the fact that other people may have beliefs which are incompatible with their own. All this goes against Piaget's view of the purely 'egocentric' child: it suggests that children are born **expecting** others to have minds, and their behaviour to be governed by mental states.

Moral/social theories

Here a variety of work has been done by, e.g. Cosmides, Tooby and Premack, studying the development of the child's innately determined moral and social assumptions. I'll give just two examples.

Cosmides (1989) argues that there is innately determined knowledge of social exchange structures, so that the child knows, for example, that if you accept help from someone then you owe them something in return. This would in turn give rise to 'cheat-detection' strategies, which have been claimed to underlie the evolutionary development of logic. (You detect that someone has cheated by noticing that they did not do what they promised to do: i.e. there is a contradiction between what they promised and what they did. To do this, you have to have the ability to spot contradictions.) It has been shown that children as young as two make moral distinctions between accidental and intentional behaviour, regarding an intentional act of disobedience as much more serious than an accidental one, for example (again, this plays a role in vocabulary acquisition, as we'll see next week).

Premack and Premack (1994) have provided evidence for a whole cascade of richer and richer moral/social modules, corresponding, perhaps, to something like Fodor's triggering hierarchy. So, for example, once you have the ability to spot agents and goals, you may see one agent as 'trying' and 'failing' to achieve something, and another as 'helping' or 'hindering' it in the attempt to achieve a goal. You may see one agent as 'reciprocating' the help or harm done to it by another. You may then distinguish 'possessions', which move along with the possessor, from 'free agents', which are self-propelled. You may be able to conceptualise 'groups' as collections of free agents who share reciprocity. You may be able to distinguish what a free agent 'ought' to do because of the expectations of reciprocity among other group members – and there is experimental evidence for all of this. In this way, the whole form of a moral theory (though not necessarily its content) is seen as resulting from a cascade of innately determined modules, each having to be triggered before the next comes online.

The number sense (Dehaene 1997; Wynn 1998)

A series of experiments have shown that infants have two types of numerical ability:

(a) An ability to keep track of up to 4 **distinct objects** (found in human infants and adults, and in rhesus monkeys)
(b) An ability to make **gross comparisons** among larger sets of objects (shared by many species).

The ability to **track objects** allows 6 month old infants to notice when an object appears in or disappears from a group of 3-4 objects (for example, they are surprised if, having seen 2 objects roll behind a screen, they find only 1 when the screen is removed). The ability to make

gross comparisons among larger sets enables 6 month old infants to distinguish a display of 4 objects from one of 8 objects, or one of 8 objects from one of 16 objects (for success, the ratio has to be something like 2 to 1). Interestingly, these abilities seem to be quite distinct. For example, although the infant can keep **track** of up to 4 objects, it is unable to **distinguish** between two displays, one containing 2 objects and one containing 4 objects: this ability to make gross comparisons only applies to larger sets (Xu 2003).

Further systems that have been investigated include **biological kinds** (Atran 1987; Dupré 1981), **emotions, geometry, artefacts**, **music** and, of course, **language.** This seems to bear out the suggestion of evolutionary psychologists that our minds are 'adaptive toolkits', made up of special-purpose **mental organs**, each with its own system of **representations** and **computations** (or **procedures**)**.**

3. Theories or modules?

There is general agreement among developmental psychologists in this very rich tradition that classical empiricism is wrong: the initial state is innately determined, or **triggered** (i.e. the child is born with a lot of expectations about the world), and different types of expectation are linked to different cognitive domains. However, there is still room for disagreement over what happens next. **Modified empiricists** (e.g. Carey, Gopnik) say that the initial endowment is a basic **theory** of a certain domain, and that what happens after this is further **rational theory construction,** based on Fodorian central processes. On this approach, the child starts with an initial theory of a domain (e.g. other minds), and approaches data from that domain like a little scientist, noticing correlations and coming up with richer and richer theories to explain them, assessing evidence for one theory and modifying it in the face of counter-evidence. The implication for concept acquisition is that after the initial impetus, later concepts are **learned** by rational processes of hypothesis formation and confirmation. **Modified nativists** (who include Leslie and Premack) say that the initial endowment is not so much a theory as a set of **data** and **procedures/mechanisms**, and that what happens after the initial triggering is not rational inference but more triggering of higher-level modules, along the lines sketched above for naive physics, mindreading, and for Premack & Premack's cascade of social/moral modules. Dan Sperber's recent work (e.g. Sperber 1994) is an attempt to work out an overall architecture along modified-nativist lines, and Gopnik 1996 sketches an overall architecture on modified empiricist lines. (For where pragmatics fits in, see Sperber 2000; Sperber & Wilson 2002, Wilson 2003.) The modified empiricists are sometimes called 'theory' theorists and the

modified nativists 'modularity' theorists, or ‘massive modularity’ theorists, depending on how many of Fodor’s central systems they see as having modular structure.

How do we choose between the two approaches? Here I'll summarise some of the arguments of Segal (1996), which point towards a modified nativist approach and away from the ‘child as scientist’ view (some of the arguments are more against classical empiricism a la Piaget, but some also work against modified empiricism):

1. On the ‘child as scientist’ approach, the child has to be not just a little scientist but a brilliant one, but children’s abilities are rather patchy (they can do naive physics and theory of mind, but do not know, e.g. the principle of inertia). This suggests special-purpose abilities rather than general-purpose ones.
2. The patterns and results of development are very similar across individuals and cultures. If children are little scientists, how do they find the same hypotheses so quickly, seemingly independently of general intelligence or teaching? This suggests special-purpose rather than general-purpose abilities. (‘Poverty of the stimulus’ argument.)
3. One of the main forces in genuine scientific change is conscious reflection: the scientist is explicitly aware of counter-evidence and consciously tries to come up with an alternative account. This is clearly not what happens with children.
4. Perhaps the most compelling argument for some version of nativism is the existence of double dissociations between general intelligence and specialised abilities. This seems to happen with mind-reading. For example, Williams Syndrome is a rare genetic disorder resulting in average IQ of around 50, but combined with a high level of linguistic ability and, crucially, relatively high social skills. Williams Syndrome children pass the False-Belief (Sally-Anne) task with ease (mindreading), and also pass tests requiring the ability to attribute goals. However, their general ability to acquire theoretical, explanatory knowledge is very poor. Similarly, as is well known, there are highly intelligent Asperger Syndrome individuals with good general explanatory skills but very deficient mindreading skills. If anything, these people show what it is like to acquire mindreading by reflection rather than by instinct, and their differences from others confirm the modularity of mindreading (Sperber & Wilson 2002)
5. Despite what 'theory theorists' say, it's hard to see how richer concepts can emerge by induction from more impoverished theories. This was Fodor’s original argument (in *Language of Thought*), which started the whole recent innatist programme in developmental psychology.

While these arguments are not conclusive (and could not be until we know what the modified empiricists regard as the innate endowment, and what they regard as acquired by

rational means), they do seem to point in a broadly nativist rather than empiricist direction.

4. Implications for the analysis of concepts

One way of looking at the modified nativist approach (particularly the massively modular version proposed by Sperber 1994, 2002), is to see it as attributing to the child an innately determined **format** for concepts in different cognitive domains. As we've seen, a concept (conceptual address) may be thought of as giving access to a **detector** (e.g. the frog's moving-black-dot mechanism for detecting flies), to encyclopaedic information (**data)** (some of which might be innately determined), and to **procedures/mechanisms**, which may be of many different types (e.g. logical inference rules, which might take different formats in different domains; attention-directing mechanisms, which might direct attention in different ways in different cognitive domains; procedures for activating information in encyclopaedic memory, which might create different patterns from domain to domain, and so on). The experimental evidence in section 3 might then be seen as providing insight into the general format for concepts of **physical objects**, **living kinds, people, faces, emotions, artefacts, numbers, language, music**, and so on, with some wired-in **data**, **inferential procedures, attention-directing mechanisms, activation patterns, etc.** Indefinitely many individual concepts with this format could then be triggered ('initialised') by experience in different domains, but the child would not be starting from scratch, with no expecations: rather, it might have quite definite expecations in different domains.

To give some (very rough) idea of how these formats might look, let's take two examples from domains that have been well-studied from both philosophical and psychological points of view: **natural kind (and particularly animal kind)** concepts, and **artefact concepts** (for tools and other items made by humans for their own use). According to Margolis (1998), who takes a broadly Fodorian (atomistic) approach to concepts, animal-kind concepts like CAT, BIRD, HORSE, etc. (and more generally natural-kind concepts like WATER, GOLD) might have the following format:

(a) A **detector** (what Margolis calls a **syndrome)** constructed by mechanisms that direct the child's attention to certain superficial properties of objects that are clearly self-propelled: e.g. shape, size, number of legs, face, covering (skin, feathers, fur), movement, etc.
(b) An **essentialist** assumption (for which there is good experimental evidence) that children expect these superficial properties to be caused by some internal structure which is common to all members of the species (or kind).

(c) The assumption that what is important to membership of the species (or kind) is not the superficial properties but the essence. So, for example, an object with the same superficial properties but different essence would not be regarded as a member of the species (or, more generally, the same natural kind) – cf. Putnam's arguments about 'water' on Earth and Twin Earth not belonging to the same natural kind because they lack the same internal causal structure), and an object with the right essence which lacks some of the superficial properties would still count as a member of the species (e.g. a horse painted to look like a zebra would be classified as a horse, not a zebra, by children who are told about the painting, cf. Carey's developmental evidence).

According to psychologists like Bloom (1996) and philosophers like Millikan (1998), artefact concepts like CHAIR, CLOCK, CAR, etc. behave differently from natural kind concepts in at least some respects. For example, artefact concepts don't have a hidden causal structure, or essence. If you notice several camels with humps, you will tend to infer that all camels have humps; whereas if you see several chairs with plastic seats, you won't tend to infer that all chairs have plastic seats. Here is the sort of format that might be proposed:

(a) We may well have **detectors** based on physical shape, but since artefacts may still perform their function with wildly different shapes, these detectors may be much more flexible and idiosyncratic than for animal kinds.
(b) Parallel to the essentialist assumption for natural kinds is the assumption that what is important to an artefact is its **function** (or, for Bloom, its **intended function,** or the function for which it was made).
(c) The assumption that what is essential to membership in the kind is the function (or intended function) rather than the superficial properties (so an art-object shaped like a boat, that wouldn't actually sail, would not be regarded as a real boat (even if it was loosely described as a 'boat'). (Bloom has a nice survey of possible approaches to artefact concepts, and many possible counterexamples to most accounts. See also Sperber 2003 'Seedless grapes' (downloadable from his website) for a fascinating discussion of the notion of artefacts, and of biological and cultural functions). (For more philosophical discussion on artefacts, see Putnam 1975, who argues that artefact terms have real essences, and Schwartz 1978 'Putnam on artifacts' *Philosophical Review 87, 566-574*; Schwartz 1980 'Natural kinds and nominal kinds'. *Mind* 89, 182-195; see also Millikan 1998, plus discussion of her article in BBS.)

Reading

Margolis, E. 1998 How to acquire a concept. *Mind and Language* 13: 347-69. (Reprinted in Margolis & Lawrence (eds) 1999 *Concepts: Core Readings.* (Particularly on *syndrome-based sustaining mechanisms*).

Homework

Make sure you understand Margolis's account of how Fodorian natural-kind concepts can be acquired.

References

Baillargeon, R. 1993. The object concept revisited. Reprinted in Margolis & Lawrence (eds) *Concepts: Core Readings.*

Bloom, P. 1996 Intention, history and artifact concepts. *Cognition* 60: 1-29.

Leslie, A. 1994 ToMM, ToBY and Agency: Core architecture and domain specificity. In L. Hirschfeld & S. Gelman (eds) *Mapping the Mind: Domain Specificity in Communication and Culture*. CUP.

Segal, G. 1996 The modularity of theory of mind. In P. Carruthers & P. Smith *Theories of Theories of Mind*. CUP.

Carey, S. 1991 Knowledge acquisition: Enrichment or conceptual change? In S. Carey & R. Gelman (eds) *The Epigenesis of Mind: Essays on Biology and Cognition.* Also in Margolis & Laurence: 458-87.

Dehaene, S. 1997 *The Number Sense.* NY: OUP.

Premack, D. & A. Premack 1994 Moral belief: Form versus content. In Hirschfeld & S. Gelman *Mapping..Mind.*.

Xu, F. 2003 Numerosity discrimination in infants: Evidence for 2 systems of representations. *Cognition* 89.1: B15-B25.

Atran, S. 1987 Ordinary constraints on the semantics of living kinds. *Mind & Language* 2: 27-60

Keil, F. 1992 *Concepts, Kinds and Cognitive Development*. MIT Press, chapter 3.

Premack, D. 1990 The infant's theory of self-propelled objects. *Cognition* 36: 1-16.

Gallistel, C. & R. Gelman 2000 Non-verbal numerical cognition: from reals to integers. *Trends in Cognitive Sciences 4: 59-65.*

Margolis, E. 1998 How to acquire a concept. *Mind and Language* 13: 347-69.

Xu, F. 2003 Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition* 89.1: B15-B25.

Atran, S. 1987 Ordinary constraints on the semantics of living kinds. *Mind & Language* 2: 27-60

Dupré 1981 Biological taxa as natural kinds *Philosophical Review* 90, 66-90.

Millikan, R. 1998 A common structure for concepts of individuals, stuffs and real kinds: More Mama, more milk, and more mouse. *Behavioral & Brain Sciences* 21: 55-100. (also in Margolis and Laurence). Interesting discussion of kinds of concepts.

Sperber, D. 1994 In Hirschfeld & Gelman (eds) *Mapping the Mind.* downloadable from www.dan.sperber.com

Sperber, D. 2002. Massive modularity. In E. Dupoux (ed). downloadable from www.dan.sperber.com