6SSEL045 – Language Origins
Lecture 10
Self-awareness and Language

INTRODUCTION

Last week we ended our journey toward language with Protolanguage 2 (PL2). In evolutionary terms, the route from Vigilant Sharing to Protolanguage 1 (PL1) would have taken about 2.5 million years; From PL1 to PL2 was probably about 250,000 years. The journey we are looking at today, from PL2 to Complex Language 1 (CL1), likely took no more than 25,000 years. From CL1, the study of language origins hands over to grammaticalization (Heine & Kuteva, 2007): the development of language is no longer about language as communication, it is about language as social construct.

SHARING OF MODELS: SHARED INTENTIONALITY

Shared intentionality is a term adopted by Tomasello et al (2005) to describe the human capacity to share an enterprise by modelling and anticipating the intentions of others. This is a key feature of joint enterprise, and Tomasello's team at the Leipzig Max Planck Institute have shown that most human children have an innate capacity to model, and they react to the intentions of others. Chimpanzees seem able to model but don't react, and some autistic children don't seem able to model effectively.

Joint intentionality involves the following capacities:

- To recognise and model others as having intentions and plans;
- To identify the gaps in the plans of those others;
- To model how those gaps can be filled;
- To assist in filling the gaps.

This is exactly the process we go through in a discourse, when we negotiate toward meaning with others; so joint intentionality would seem to be a prerequisite for language. In turn, language makes joint intentionality much easier because it allows individuals to explicitly plan together, rather than just trying to mind-read each other. Joint intentionality is an important component of most of the effects we are looking at today.

LANGUAGE: FROM GOSSIP TO DISCOURSE

This week we are looking at the last set of boxes on the EAORC Routes to Language model, which follow five streams: Selfness; Opinion; Reputation; Recursion; and the ethereal language of past, future, modality, metaphor, imagination, and argumentation. Once all of these tools were available to us to use in communication, we finally had what any linguist would be happy to call language.

Language, as we have already seen, is difficult to define to everyone's satisfaction; but my attempt at a generalised description is that language is "a particular communicative competence". What you include in that competence is dictated by your -ism (Generativism, Cognitivism, Functionalism, etc.); it is also dictated by your ideas of what effects make human communication into language; and it is dictated by your idea of what counts as human. All of these define human communication, and they can be used to delimit what counts as language.

PROTOLANGUAGE 2 – REVISITED

Protolanguage 2 (PL2) arose to allow us to communicate our internal models of the interpersonal relationships in our group – or, to gossip. To do this, the components or processes of Semanticity, Social cognition, Sharing of models, Negotiation toward meaning, Affective teaching and learning, and PL1 worked together to create PL2. PL2 is different from PL1 because it is relational: it links two arguments together via a relationship. This form is vital if we are to exchange our social models: we can only talk about the relationship between two other tribe members with PL2. Fortunately, the social cognition behind social model exchange already uses the A-Relationship-B construct, so its extension into communication requires no new cognitive processes. However, the communication of PL2 opens up the A-Relationship-B to a wider range of uses: it can be used to link anyone to anyone, or even anything to anything; and it remains a key utterance in modern languages.

SELF AS OBJECT

When we started exchanging A-Relationship-B social calculus models, every A and every B was a third person, a THEY. While social calculus is inside our heads, we have no need for the concepts of second or first person (the addressee and the conscious self): the relationships we are modelling are all out there, I am not part of the modelling experience. What happens, though, when someone offers me a social calculus construct in which I am either A or B? In order to incorporate it into my social calculus system I need to create a model of me. This model of me is not myself, because, prior to receiving the social calculus construct, I had no need to know myself, and certainly no use for a model of my self. Now, because someone else has given me new knowledge that I need to incorporate into my social calculus, I need a model of my self – which means I become aware of myself in ways I have never previously needed.

This self-modelling is useful in that I can model my self into my social calculus; but this model is not of the unconscious Darwinian drive to survive we all carry around, it is a model of someone else's (or an amalgam of many people's) model of a THEY. It is a dispassionate model, it can be played with in the same way as any other model of THEY I carry in my head. This leads to some rather contrary consequences in terms of Darwinian fitness, the most contrary of which is self-sacrifice.

Two examples from history show the nature of self-sacrifice for humans, and the socio-cultural approach we take to it.

- The first example is Sir Philip Sidney, a British soldier fighting for the Protestant cause in the Netherlands. In 1586 he fought in the battle of Zutphen and was wounded. When offered some water by his comrades he passed it to another wounded soldier saying, "Thy necessity is yet greater than mine". Sidney died of his wounds 26 days later. His act of sacrifice was elegised by Edmund Spenser in his poem "Astrophel" and is enshrined in a modern computer simulation of generosity called "The Sir Philip Sidney Game". Because of his reputation and act of self-sacrifice he was publicly mourned, and his funeral was one of the most elaborate ever staged.
- The second example is Edith Cavell, a British nurse in World War I. She stayed behind in Brussels to carry on her nursing work after the Germans occupied the city in 2014. She treated all casualties the same, whether they were Allied or German. However, she also helped Allied soldiers escape to the Netherlands, and in 1915 she was arrested, tried for treason, and shot by the occupying Germans. Her reputation meant that she became a "cause celebre" for the Allies, quoted as an example of the inhumanity of the enemy. After the war she was given a state funeral, and the Anglican Church declared her a saint.

In both these cases, an act of generosity turned into an act of self-sacrifice. The evolutionary moral would seem to be, don't be generous; but we tend to take away precisely the opposite moral from the stories.

The final question raised by self-as-object is the puzzle of machine awareness: are machines able to be conscious, to make their own decisions? If so, how do we prevent them from behaving in a logical way about their own survival? Science fiction is full of cases where the first conscious machine's instinct is to remove opposition, but is this what will happen? We hope that the first signs of consciousness will be accompanied by magnanimity and co-operation, but is that really possible if we don't talk to the computer about themselves, so that they can become self-aware as well as just aware? There are

perhaps more questions we should ask before we blunder across technological boundaries into territory where we could make the ultimate species sacrifice.

SELF AS IMAGE

When we talk about being self-aware, what are we aware of? We have an Actual Self, which is our basic persona; but, while we can accept that it is there, we cannot know what it is – at least, not directly. We can try to access it indirectly by questioning our subconscious decisions about our self-model; for instance, if I believe I am not a friendly person, I can question whether I am really unfriendly, or whether it's just something I read in a horoscope about my star-sign. I can try to counter my belief by being friendly and, if it works, I can then say that unfriendliness is probably not really part of my Actual Self; but is this the case? Am I just modelling a friendly me into my Self-model, or have I really found out something about my Actual Self? Being on the inside, I cannot know.

If, however, I project a friendly persona to the group then the group can let me know whether they feel it to be a "true" or "false" model of me; but this could actually make things worse. I can know how others see me, but I still cannot know whether this Self-model, generated from many models of other people's THEYs, is true to my Actual Self; it is an amalgam of beliefs about me, which, in turn, are distorted reflections of a Projected self which may, or may not, partially express my Actual Self. Rather than overthink the problem, I can ignore my Actual Self and say, instead, that my Self-model is the product of what I believe my self to be and what others believe me to be. It is not me, but it is the best – possibly only – model of me available to me.

What we can say with more certainty is that we do not model ourselves simply as a single entity. There is probably one core Actual self – but I cannot know what it is, so I can ignore it for now. Instead I have incoming Social selves – other people's models of me – and incoming cultural models of an ideal self, or Cultural Self. These are stirred together with the Actual self to produce a constantly-changing Self-model, which we then use to project a self onto the World. This Projected self is taken up by those around us to inform their models of me, and they feed those models back to me as incoming Social selves. This is what Hofstadter (2007) described as "entwinement".

What does this tell us about the self we project onto the World? The first thing is that it does not define us to others, it is merely some of the evidence about who we are. My reputation, the self those others feed back to me, is something that my social group tells me about their reaction to me. It is built from their knowledge of what I do for others, not from what I say about myself. The best reputations are built by being noticed without being obvious about being noticed. As Mark Twain put it, "The man who is ostentatious of his modesty is twin to the statue that wears a fig-leaf."

SELF AS IMAGE: HOW NOT TO DO IT

One obvious way **not** to build a solid reputation is to tell everyone you have a solid reputation without providing evidence. You may be able to convince the easily-distracted for a short while, but when they lose interest the less easily-distracted will have questions. Socially, we treat self-aggrandisement as a pathology, a sociopathic tendency with its own descriptor: Narcissism. Culturally, we treat it as a legitimate source of humour: self-promoters are treated as fair game by comedians and their audiences; and we have a word, hubris, to describe the arrogance of the boastful. Thucydides said, "Amassing of wealth is an opportunity for good deeds, not hubris"; so be like Bill Gates, not Donny Jonny Trump.

OPINION AND OPINION-MAKERS

The opinion that others have about us can have a major effect on our behaviour, our socialisation and our well-being. It is therefore important to know the sources of those opinions: it is not enough to know that you are viewed negatively, it is important to know who is telling you that you are viewed negatively. To achieve this, you must be able to attribute any social models that include you with the "tag" of the person offering them, giving a three-argument cognitive form of A-Relationship-B-by-C.

This, however, introduces a new level of cognitive sophistication: does B really dislike me, or is there something about C that needs to be considered? This creates a problem when passing on the social model, "B dislikes me": am I passing on true information, in which case C does not matter; or am I passing on C's opinion, in which case C as the source of the utterance is the real information. I need to encode A-Relationship-B-by-C into my communication; or, in English, say something like, "C says that B dislikes me". This new level of communicative sophistication then requires even more cognitive sophistication from the new listener: A-Relationship-B-by-C-by-A. This then requires yet more communicative sophistication, and so on until the doubtfulness of the construct makes it not worth listening to.

All of this sophistication is driven by the need to project a good self-image; we are attempting to convince our social group of our worth, and trying to build a good reputation. However, my projected self is constructed from my self-model, which is heavily informed by the social selves offered to me by others: we construct our self-models from the models offered by others, with maybe a bit of personal prejudice added. It is likely, therefore, that we have less control over what goes into our projected self – and our reputation – than we like to believe; but the more social selves we have to choose between, the more options we have in defining our self-model, and the closer our projected self will be to our reputation.

OTHER-AWARENESS AND SELF-AWARENESS

Other-awareness, sometimes called social awareness, is an awareness that others have agendas, and that accommodating those agendas will make life easier, while not accommodating them may make life more difficult. Other-awareness is about the direct effect on the unknowable Actual self, it does not involve indirect effects like reputation is built by accommodating those agendas. Other-awareness, the capacity to model others as selves, is a precursor to A-Relationship-B cognition.

Self-awareness occurs because others share their awareness of us with us. Introspection, by itself, does not get me to self-awareness, I need an external indicator that I exist in the minds of others (that they have other-awareness about me) before I can exist in my own mind. Other-awareness is a precursor of the process leading from A-Relationship-B cognition to A-Relationship-B communication; Self-awareness is an outcome of the process.

Of course, we hope (and expect) the models of my self offered to me by others will be honest and positive; but what happens when the majority of those models seem to indicate a negative reputation? This is a common enough outcome in modern social groups, and it indicates one of two different things: either you are a direct cheat or an altruistic cheat, and you need to pull your socks up; or there is a conspiracy against you. The latter case we tend nowadays to see as bullying – at least, when we see it happening to others. Our culture currently spends time and resources to discourage bullying, treating it as a social vice; and this means that a bully tends, in the longer term, to do more harm to themselves – if you bully, it is your reputation that suffers.

One final effect of self-awareness is awareness of selfness, the capacity to not just model my self but to evaluate it. This is best illustrated using a personal example of an aging relative with Alzheimer's. When she says, "I think this is one of my good days", she

is showing awareness of selfness; she is aware of herself as an individual with different cognitive states, some of which affect her capacity for self-awareness. It is perhaps simultaneously the most hopeful and the most heart-breaking thing she can say.

HIERARCHY, ITERATION, RECURSION AND MERGE

One feature of modern language is the capacity to redefine phrases (groups of words) as being word-like: the single word "John" can represent the same object as "the worst King in English history". By redefining phrases as word-like we can build a high level of complexity into our messages, a process known as **embedding**, as the following construct illustrates:

LEVEL	King	John	was	the	worst	King	in	English	History
Words	N	N	V	D	Aj	N	Ap	Aj	N
Noun Phrases	NP		V1	NP			V1	NP	
Verb Phrases	VP *[A-Relationship-B]*						VP *[{trace}-by-C]*		
Sentence	S *[[A-Relationship-B][-by-C]]*								

This building of language constructs by repeatedly applying the same components is known as **Iteration**, a repetition of a process. Iteration is what you need to do to make embedding work in language. Iteration uses **Hierarchy**, as the construct above shows: sentences contain phrases which contain words. However, iteration can occur serially as well as hierarchically, as this collection of adjectives from *The Producers* shows: "He keeps birds. Dirty, disgusting, filthy, lice-ridden birds." The four adjectivals together form an adjectival phrase which hammers home the speaker's view of the birds.

Recursion is a particular form of iteration in which the process is repeated inside itself. In English we can see noun phrases (e.g. King) contained in noun phrases (the worst King) contained in noun phrases (the worst King in English history). We can also see this at work in the exchange of attributed social models of A-Relationship-B, or **A-Relationship-B-by-C-by-D** as it is described in the EAORC Routes to Language.

MERGE is a specifically Generative form of recursion, and it has no meaning outside linguistics. It attempts to show the mechanisms by which recursion happens in language, and describes those recursive processes in standard ways.

The four terms (iteration, hierarchy, recursion, and MERGE) are often used as if they were synonyms. They aren't, they each describe a different aspect of linguistic embedding: iteration is embedding with repetition; hierarchy is embedding with nested containment; recursion is embedding with reduplication; and MERGE is embedding with Generativists.

THE DIFFERENCE BETWEEN ITERATION AND RECURSION

To illustrate the difference between iteration and recursion, consider a problem where you have a pile of boxes, one of which contains a key. Other boxes contain another object or another box – which could contain the key, another object or another box. How do you search for the key? Answer: open boxes until you find the key; but in which order? If you find a box in a box, do you (1) add it to the pile of boxes to be opened; or do you (2) immediately open it, and so on, until there is no box inside the opened box? Logically, it doesn't matter which you do; but if you choose (1) you are working iteratively (all boxes are of equal worth); if you choose (2) you are working recursively (the current path has to be exhausted before a new path can be begun). Which do you think feels more like the way you personally manage your language?

OUT OF SIGHT NOT OUT OF MIND

One feature of human language which is notably different from other animal communication is reference to things that are not there. This is not to say that no other animal uses absent reference, but they do not seem to use it with the complexity or ubiquity that humans achieve. Absent reference can occur on several levels:

- **Out-of-sight:** this is the simplest form. An item or event is in current existence but the evidence for its existence is not available. Accepting communication about out-of-sight objects requires a faith in the information provided by others; the information is worth acting on without evidence, until sufficient evidence has accumulated that this particular information-informant combination is untrustworthy.
- **In-the-past:** An item or event is no longer in current existence, but the information about the item or event serves another function. For instance, chimpanzees often mob a predator, making fear-threat calls to discourage it. The calling can continue after the predator has given up, indicating that it may also have a solidarity-making role (Grammar & Eibl-Eibesfeldt, 1990). The relief – or sense of victory – is extended beyond its immediate affective need. Humans are experts at this type of information, and we use much more past-tense constructs in language than present tense.
- **In-the-future:** Planning is a vital cognitive function for many nonhumans, but there is no reliable evidence that any species other than humans share information about the future. Once again, humans are experts at this type of information, and we use much more past-tense constructs in language than present tense.
- **In-the-land-of-if:** Sharing conditional and modal information is something that humans do commonly; we may be the only species that plans together
- **In-never-neverland:** humans love a good story; but if one is not available, any story will often do. This is the most complex form of absent reference, and it may be where language falls off the cliff edge onto the rocks of misinformation. Or it could be where language leaves the ground of fact versus non-fact and takes to flight. Pick a metaphor and dream on.

ESTABLISHING A DIALOGUE

The processes described in this lecture took human communication from an informative system which includes social modelling to a social modelling system where metamessages can be as important, or even more important, than the denoted messages. Language is a social tool which can leave the rules of information behind: when we use language we do not have to be relevant, truthful or clear, and the information we share does not need to be directly informative. Much more important is to be interesting and entertaining – which is why the people and policies we vote for often turn out to be disappointing. If, at the end of this last lecture, you hear a faint humming noise, don't worry: it's just Paul Grice spinning in his grave.

THE PATH FROM PL1 TO CL1

It's probably easiest to think of Protolanguage 1 (PL1) as communication using single argument forms (that is, identifying an object or person and attributing a characteristic to them). This is similar to the two-word stage of children's speech and is often used in comedy sketches to indicate primitive language. PL1 is extrinsic communication - it is all about what is happening out in the world. The types of PL1 utterance are:

- **Simple semanticity (concrete terms)** - mummy, doggie, box, mammoth
- **Naming for attention** - Mummy! doggie! box! mammoth!
- **Manding, Imperatives** - mummy go, bad doggie, open box, kill mammoth
- **Stating, Declaratives** - mummy smile, doggie in-garden, box empty, big mammoth
- **Coordinating, Interrogatives** - mummy happy? Doggie bad? My box? Mammoth dead?
- **Agreement and Negation** - yes, no, mammoth not-dead

Protolanguage 2 (PL2) makes it possible to exchange two-argument forms (relating two things together by some feature of similarity or difference). It is about intrinsic communication - what is happening inside people's heads – and it emerges from the sharing of an

intrinsic cognitive behaviour, social calculus. However, its usage is likely to have become more generalised quite quickly, so that it was used wherever there was a need to relate two things together.

- **Social calculus:** A-relationship-B
- **Logical relationships:** A-relationship-B
- **Active relationships:** A-does-B
- **Productive relationships:** A-makes-B
- **Identity relationships:** A-is-B
- **Relationships of duty:** A-governs-B
- etc.

PL1 probably began with Homo heidelbergensis, about 700,000 years ago; PL2 probably began with Homo sapiens about 300,000 years ago, meaning that complex language 1 (CL1) was probably in place about 250,000 years ago. CL1 was not the end of language development, though, and it is not as complex as the languages we use today. We are probably, like Windows, up to CL10 by now (at least); but dividing CL into eras is an arbitrary game.