

# A new kind of language

Paul Ibbotson discusses an emerging radical view of language acquisition

**If there ever was a time when language was seen as a medium we could use without paying too much attention to it, then that time is over – we no longer take language for granted. The medium has become the object of attention, so much so that the exposition of language has come to be one of the defining intellectual characteristic of our age. Part of this inquiry wants to understand how we learn language in the first place. Learning to speak is such a ubiquitous feature of childhood it seems an inevitable and – perhaps on the face of it – a trivial aspect of human development. But...**

As the cognitive scientist Marvin Minsky recognised, 'we're least aware of what our minds do best, accordingly, we're more aware of simple processes that don't work well than of complex ones that work flawlessly' (Minsky, 1987, p.29). By studying language in more and more detail over the last few decades we have been able to raise our awareness of how powerful, complex and subtle it really is. In fact, everything we know about language suggests a system for thinking and communication that is flexible and intricately structured. We also know that children show development towards adult mastery over this system. But how does this happen?

For a long time, the leading contender for a theory of how children master their language was proposed by Noam Chomsky. In a top 10 of the most cited, Chomsky is just behind Freud, Plato and the Bible and the only living member on the list. His most famous brainchild is universal grammar – the design underlying all human languages, past, present and future – which is coded in our genes and which allows children, but not our closest evolutionary cousins, the chimpanzees, to acquire their mother tongue. In the story of Babel, God scattered humanity about the face of the earth, confounding their language so that they may not understand one another's speech. Chomsky's seductive idea was to reunify mankind at a deeper level; despite the surface differences in Italian, Gujarati, Yiddish and Navajo they are all variations on an underlying theme: universal grammar. But does it work?

Chomsky defined universal grammar in terms of mental software that was standard issue for every child. This combined two big ideas which were right for the times – computational theory of mind and behavioural genetics. Claude Shannon's information-theoretic approach to communication deliberately abstracted away from any meaning in the message. It was in this spirit that Chomsky defined grammar. Like computer software, the principles of grammar would follow a script, cranking through algebraic rules until it was told to stop. Also like computer software, the operations themselves are meaningless to the system that carries them out. By applying production rules to sequences of symbols, a formal grammar could generate an infinite set of finite-length sequences of symbols. The mathematisation of language required Chomsky's idealised speaker-listeners to exist in a 'completely homogenous speech-community... unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors in applying his knowledge of the language in actual performance' (Chomsky, 1965, pp.3–4). Chomsky's proposal also resonated with another intellectual trend. Evolutionary psychology was establishing itself as a subfield of evolutionary biology. Universal grammar made strong claims about heredity that, if true, would mean that the phenotypic variation we see in the world's languages is a product of a universal grammar genotype.

It's over 40 years since Chomsky first proposed this idea, and since then it has inspired tens of thousands of scientific papers dedicated to uncovering what the contents of the universal grammar might look like. Despite this sustained effort, when researchers actually look across the 6000-plus languages in the world for traces of universals, what they find instead is diversity: some languages make words from other words, like 'run' to 'runner', others don't; some languages have notions for 'left of', 'right of', 'back of', 'front of',

## questions

How do children move from single word expressions to more complex linguistic forms?

What explains the similarities and the differences between languages?

What allows humans but not chimpanzees to learn language? (Despite our best efforts to teach them.)

## resources

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Free online corpus of child speech and child-directed speech in over 20 different languages:  
<http://childes.psy.cmu.edu>

## references

Ambridge, B., Rowland, Theakston, A.L. & Tomasello, M. (2006). Comparing different accounts of inversion errors in children's non-subject wh-questions. *Journal of Child Language*, 33, 519–557.

Cameron-Faultkner, T., Lieven, E. and Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science*, 27, 843–873.

Carpenter, M., Nagell, K. & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63 (4, Serial No. 255)

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA, MIT Press.

Evans, N. & Levinson, S. (2009). The myth of language universals. *Behavioural*

*and Brain Sciences*, 32, 429–492.

Inagaki, K. & Hatano, G. (1987). Young children's spontaneous personification as analogy. *Child Development*, 58, 1013–1020.

Minsky, M. (1987). *The Society of Mind*. Heinemann. London.

Rowland, C.F. & Pine, J.M. (2000). Subject-auxiliary inversion errors and wh-question acquisition. *Journal of Child Language*, 27, 157–181.

others don't; some languages have adverbs and adjectives, others don't; some have a fixed order of items others don't; some have recursion, others don't. For example, in English the sentence 'John likes Mary' can be embedded within 'Susan thinks John likes Mary' or embedded further within 'I suspect Susan thinks John likes Mary', and so on. In theory, this fractal-like property means infinite productivity from a finite list of rules. Yet some languages, such as the Amazonian language Pirahã, do not seem to accommodate sentence embedding and other forms of recursion (Evans & Levinson, 2009).

The principles and parameters theory was an attempt to deal with this cross-linguistic variance. It said all languages operate by means of similar principles, but with parametric variations. For example, there might be a parameter, call it 'P1', that allows a speaker to omit the subject in a tensed sentence with an inflected verb.



The new view of language sees diversity and similarity as two sides of the same coin

English children would set the parameter to 'off' because they don't hear sentences of the type 'goes to the shop' whereas Italian children do hear the equivalent in their language and would therefore set it to 'on'.

An important part of the proposal was that once a parameter was set, it would cause changes to cascade through the grammar. This elevated the setting of a parameter from merely learning an isolated fact about their language, to predicting a variety of more subtle details. Setting P1 'on' would mean that the child also 'knows' it can say things like 'Who do you think that left?' to mean 'who left?', which is fine in Italian but not in English. So, universal grammar had all of the possible parametric variations in it, and children just had to determine early in their language development what kind of language they were hearing.

The problem was, and still is, that it is

difficult to describe principles, even at a quite general and abstract level, that would help the child in learning. This is because any general principle (even 'subject' and 'noun') seems to work differently in different languages – the cascade doesn't get very far, if it starts at all, before it runs afoul of the cross-linguistic facts. In the above example, the hypothesised parameter P1 is predicated on the idea that all languages will have more-or-less the same basic grammatical category of subject. The problem is that 'subject' looks different in different languages, so in order to capture the diversity, the definition of subject becomes so vague that it offers little help to a child in parameter setting. In theory one could keep adding more microparameters that do capture the variation, but this dilutes the original principles and parameters idea to such an extent that parameter setting begins to look indistinguishable from learning all the idiosyncratic facts about a language.

Scientists are traditionally bad at abandoning an idea, even when it is struggling, unless there is a ready alternative available. As a radically different view of language has started to gain respectability, many cognitive scientists and linguists have been gradually turning their back on the idea of universal grammar. The new view has consequences

for how we think children learn language, and if these emerging ideas turn out to be right, we need to think again about how

language works, how it evolved and about human uniqueness. We turn now to some key elements of the new view.

### Language as history

Despite the diversity across the world's languages, there are regularities too: the question is how best to explain them. Those defending the idea of universal grammar, such as Steven Pinker, point to the fact that there are many languages that could exist within the bounds of human cognition but don't, and potentially, universal grammar can explain why. The new view of language sees diversity and similarity as two sides of the same coin: products of biocultural evolution. Evolution has to work with what it's got, cumulatively tinkering with solutions to problems that have worked well enough for previous generations, but

which might not be considered 'ideal' if one could start afresh (e.g. the backward installed retina in humans). Likewise, languages represent a cumulative record, and recent analyses of many languages have shown that cross-linguistic variation is best explained as a product of cultural-historical factors and constraints of human cognition.

Here is an analogy. An evolutionary biologist could ask the question 'Why do we only see the mammals we do given all the possible mammals that could exist but don't?' For the biologist, different mammals could be thought of as a record of competing motivations, that have over time explored some of the space of what is physiological plausible for a mammal to be. For a particular feature of the animal's development, for example the skeleton, a bat could be thought of as occupying one corner of this space while an elephant skeleton is in another – extreme variations on an underlying theme. For the linguist, different languages are also a history of competing motivations that have explored some of the space of what is communicatively possible. For a particular feature of the language development, for example the sound system, in one corner there might be a three-vowel system (e.g. Greenlandic, an Eskimo-Aleut language) while in another corner is a language with 24 vowels (e.g. !Xu, a Khoisan language). In case these examples strike English-speakers as exotic, English allows a syllable structure of consonants and vowels in words like *strengths*, which strikes speakers of many languages as very unusual. We don't need a universal grammar to explain why all spoken languages have a vowel system of some description – it's a trivial fact about human cognition that without vowels speech would be barely audible.

Evolving towards an extreme in one direction will have consequences for the system as a whole. Just as a bat skeleton will place certain functional demands on the rest of its physiology, so it is with language. In languages with freer word order, the communicative work that is done by a fixed word order in other languages must be picked up by other aspects of the system, e.g. morphology and pragmatic inference. Over time languages have adaptively radiated to different points from a similar starting point. So we suspect that Pinker's objections can be met if we take the historical dimension of languages into account: the languages we do and don't see are best explained in terms of a record of competing motivations that have partially explored what is communicatively possible.

The analogy with biological evolution

## language development

might be useful in another way. The eye has independently evolved in several different species, converging on a similar solution to a similar engineering problem. The major nuts and bolts of grammar (e.g. word-order, case-agreement, tense-aspect) that appear time after time in the world's languages could be thought of as historically popular solutions to similar communicative and coordination problems, such as sharing, requesting and informing. It is these kinds of explanations – a history of complex interactions between cognition and culture – that may ultimately explain why we see some of the tendencies across languages, once thought to reflect an underlying universal grammar.

### Language as mind reading

When the writer J.M. Coetzee was asked to name his favourite novel he replied that it was Daniel Defoe's *Robinson Crusoe*, because, he explained, the story of a man alone on an island is the only story there is. The island Coetzee refers to here is, of course, a powerful metaphor for the mind. 'I can understand 'you' by analogy to myself, but never as yourself.

Although we can never escape our own island, we still develop a strong sense that we can read other people's minds. Suppose you see a woman take keys out of her pocket while approaching the front door of a house. Not only do you perceive the physics of the scene – the solidity, acceleration and trajectory of the objects – you habitually think of mental motivations for her actions – 'she wants to open the door', 'she's trying to get in her house', 'she believes the key will open the door'. If she knows you are watching the scene she will have beliefs about what you believe of her 'you think I'm trying to get in the house', and so on. So, fundamental to interpreting human behaviour is our capacity to think about other people's actions in terms of their underlying goals, intentions and beliefs. As well as the ability to read behaviour in this way, we seem to possess an intrinsic, and perhaps species-specific, motivation to share psychological states (and know that we share these states) with other people. There are even those who argue that it is this kind of thinking that leads to a kind of consciousness, which happens when the model of the world is so detailed and flexible that it includes a model of itself.

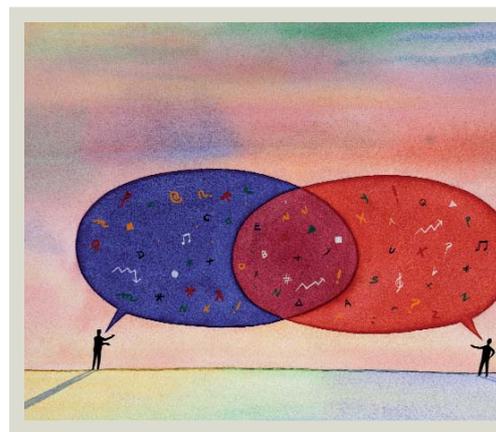
What's all this got to do with language? The central idea here is that language is a code that rests on a deeper code about how people work. We understand the mundane utterance 'Can you pass the salt?' as a transparent request rather than, say, as an enquiry about condiment passing

abilities, because we can work out what is relevant from their perspective. One way to work out what's relevant for them is to have a model in your head of their beliefs, desires and attitudes. Research in infants suggests that language begins to take off shortly after infants show signs of this mindreading ability (e.g. Carpenter et al., 1998). This is because in order to understand that language isn't just like any other string of sounds infants have to work out that it can be used to manipulate people's attention and thoughts towards an object, an idea, an attitude, and so on. Once this deeper code is understood the next code can be learnt – what sounds does this particular language use to direct people's attention.

Learning language, then, is seen as part of a broader adaptation to cultural learning. In the old view of language, the paper-and-pen analysis of grammatical patterns led linguists to conclude that there was a problem for the child; there was not enough data in the language to get where you needed to be as an adult speaker (part of why Chomsky suggested universal grammar in the first place). Researchers are starting to realise that this isn't the same problem children are faced with as social agents grounded in a communicative context.

### Language as cognition

Chomsky once wrote: 'I think a linguist can do perfectly good work in generative grammar without ever caring about questions of physical realism or what his work has to do with the structure of the mind.' The new view of language thinks differently. Researchers have begun to uncover what general cognitive processes can teach us about language, what general learning mechanisms can teach us about language acquisition, and how language is grounded in the world of the speaker. For a start, research suggests that language emerged relatively recently in human history, and therefore must have exploited (and continues to exploit) pre-existing brain machinery. Take categorisation. Humans are prodigious at finding patterns, organising experience into clusters of things that look the same or do the same thing. This includes everything from the relatively abstract generalisations of grammar to the more prosaic: when a six-year-old realises that tulips must need water, because people do. In this example, the two things being compared fall under a category of things that share a property – living thing needing water. Underscoring the 'development' in 'developmental psychology', the same



study reported that 40 per cent of toddlers believe that a tulip can 'feel happy' and 72 per cent believe that a tulip can 'feel pretty' (Inagaki & Hatano, 1987). In this example the children had overgeneralised a pattern – all things that feel are probably alive but not all things alive feel – and we see this in language as well, for example when English-speaking children say sheeps, feets and mouses.

In cognitive science, finding the common relational structure of complex entities – that is, similarities not of the items involved but of the relations among items – is called analogy, and we are beginning to understand how crucial it is to language learning. When the learner is trying to comprehend the two sentences 'the goat ate the woman' and 'a woman tickled a goat', they do not begin by aligning elements on the basis of the literal similarity between the two goats, but match the goat and the woman because they are both construed as playing similar roles in the event, such as actor or undergoer. There is much evidence that people, including young children, focus on these kinds of relations in making analogies across linguistic constructions. Analogy also functions at the level of words and sounds. When a learner says 'I goed to shops' it is by analogy to a set of verbs that mark past tense in English with —ed. Likewise, 'I brung it' is produced by analogy to 'phonological neighbours' that show similar sound alternations, such as 'sing-sung'.

To make such analogies one might argue the child is going to need a lot of raw data to work with. One analysis of child-directed speech showed that infants are spoken to on average 7000 times a day (Cameron-Faulkner et al., 2003). For these children at least, this means that between the ages of two and five, a period where language significantly develops, they hear something in the order of seven-and-a-half million utterances. This kind of research emphasises that the new view of language takes seriously how people actually use language, and it has revealed some

interesting findings. A large proportion of the utterances children hear can be accounted for by a relatively small number of types of utterance.

Utterances like 'it's a dog', 'it's a washing machine', 'it's a surprise', unfold in a semi-predictable way, a pattern that could be represented as 'it's a X'. These give children a foothold to understand and produce language in a more sophisticated way. More flexible language emerges when a number of 'X likes X', 'X wants X', 'X sees X' are recognised as instances of a more general syntactic pattern in which the linguistic items play similar communicate roles. Thus, theories of learning have begun to seriously consider the possibility that it might not be beyond the formidable processing power of 50–100 billion neurons with 100 trillion synaptic connections to detect such patterns in this distribution, especially when those patterns have social cognitive pay-offs in being able to detect them – understanding and being understood.

A good example of this pattern-detection in action is question formation in English (see box).

## Question formation in English

In generativist approaches, when a child forms a question – say, 'What is she tickling?' – they do so by following a set of rules operating on abstract grammatical categories. The child derives a question from a simple declarative phrase by cranking through a set of transformations in a stepwise manner, which might look something like this: She is tickling the goat → She is tickling what → What she is tickling → What is she tickling? The basic prediction of this is that when children start learning a language they will do so with abstract structures from the beginning, at least with regard to the particular parameter. This means that they should generalise the ability to form questions quickly and across-the-board.

Yet English-speaking children go through a phase of asking questions conventionally ('What is she tickling?') at the same time as asking them unconventionally ('Why he can't come?'). If the rule being used is about question formation in general why does it vary from question to question? This is more easily explained if we assume that the environment is not just the trigger for underlying representations proposed by generative linguists, but provides the raw materials out of which young children construct their linguistic inventories. When we look at the input children are exposed to in corpus studies (e.g. Rowland & Pine, 2000) we find that children produce questions correctly when they have heard highly frequent exemplars of the particular sequence in the input (e.g. Why is she [X]ing Y?) and they make more errors on sequences of lower frequency. In an experimental follow-up study (Ambridge et al., 2006) it was also found that children's error in questions varies by the lexical subtype of the auxiliary part of the question (*be*, *can* and *do*). Recall that generativist theory describes grammatical rules operating on abstract categorical variables such as subject, verb, object, auxiliary. This means generative grammar cannot 'see' inside these categories, to subtypes of auxiliary for example, but it appears that children can, and do make use of fine-grained lexical differences when constructing their grammar from the bottom up.

### Language as symbolic

On the generative view, all grammar is completely regular with all arbitrary and idiosyncratic phenomena belonging to the lexicon. One problem with suggesting a fundamental division of labour in language between words and rules is that there are phrases that are grammatically irregular yet productive, like 'Him be a pilot!?', 'You sing an Opera!?', 'My mother-in-law a saint?!' One way in which to deal with these types of phrases is to represent them all as constructions – symbolic assemblies of form and meaning.

The idea of a symbolic construction can be illustrated with an old joke. A stranger is at a dinner party where the other people all know each other well. Because they always tell the same jokes to each other they have decided to save time by numbering them. One of them says 'Number 46' and everybody laughs. Another one says 'Number 16' and everyone laughs again. The stranger tries his luck and says 'Number 9', which is greeted by a rather embarrassed silence. 'What's wrong?' said the stranger, 'isn't Number 9 a joke?' 'It is,' replied another guest, 'but you're not telling it very well.' 'Number X' in this case is a construction, a symbolic assembly of form and meaning. The key insight was to realise that this continuum could be extended to cover all language, not just idioms. If children can

learn idiosyncratic constructions like these, then why can they not learn the more canonical ones in the same way? Why do they need universal grammar? Construction grammar argues that the difference between the most idiomatic (e.g. *by and large*, 'Number X') to the most systematic (e.g. *John kissed Mary*) is one of degree not kind. Moreover, constructions vary not just in the extent to which they specify lexical material. They also vary in how specific or schematic their pragmatic, semantic, phonological, syntactic and discourse requirements are. Grammar isn't just a list of constructions though, these are organised into families of related constructions that share functions or form with other constructions. Grammar is then more broadly conceived as that set of resources used for communicating these constructions with other persons in particular usage events.

### Conclusions

New discoveries in linguistics, evolutionary anthropology and psychology have led many researchers to question the way Chomsky and his followers have thought about language and its acquisition. The new view of language doesn't deny something is innate which sets us apart from chimpanzees: the debate centres over what that

something is and how it enables the two-year-old to learn language.

This new synthesis is far from a consensus in language research, and the challenge to the new view of language is to integrate our species-specific social cognition with a prolific pattern-finding abilities in a way that can help the infant in learning language. If it turns out Chomsky was indeed wrong, he was wrong in an interesting way, interesting enough to inspire a whole academic industry devoted to showing how he was wrong and in doing so getting us closer to some important discoveries. The emerging picture, in its own way, is perhaps no less radical than that of Chomsky's. Instead of grammar being an innate endowment, language emerges in humans as the result of a handful of cognitive skills (e.g. pattern-finding, attention and memory) interacting in complex ways with a species-specific set of social skills (e.g. shared intentionality and cultural intelligence).



**Paul Ibbotson**  
is at the Max Planck Child Study Centre, School of Psychological Sciences, University of Manchester  
paul.ibbotson@manchester.ac.uk