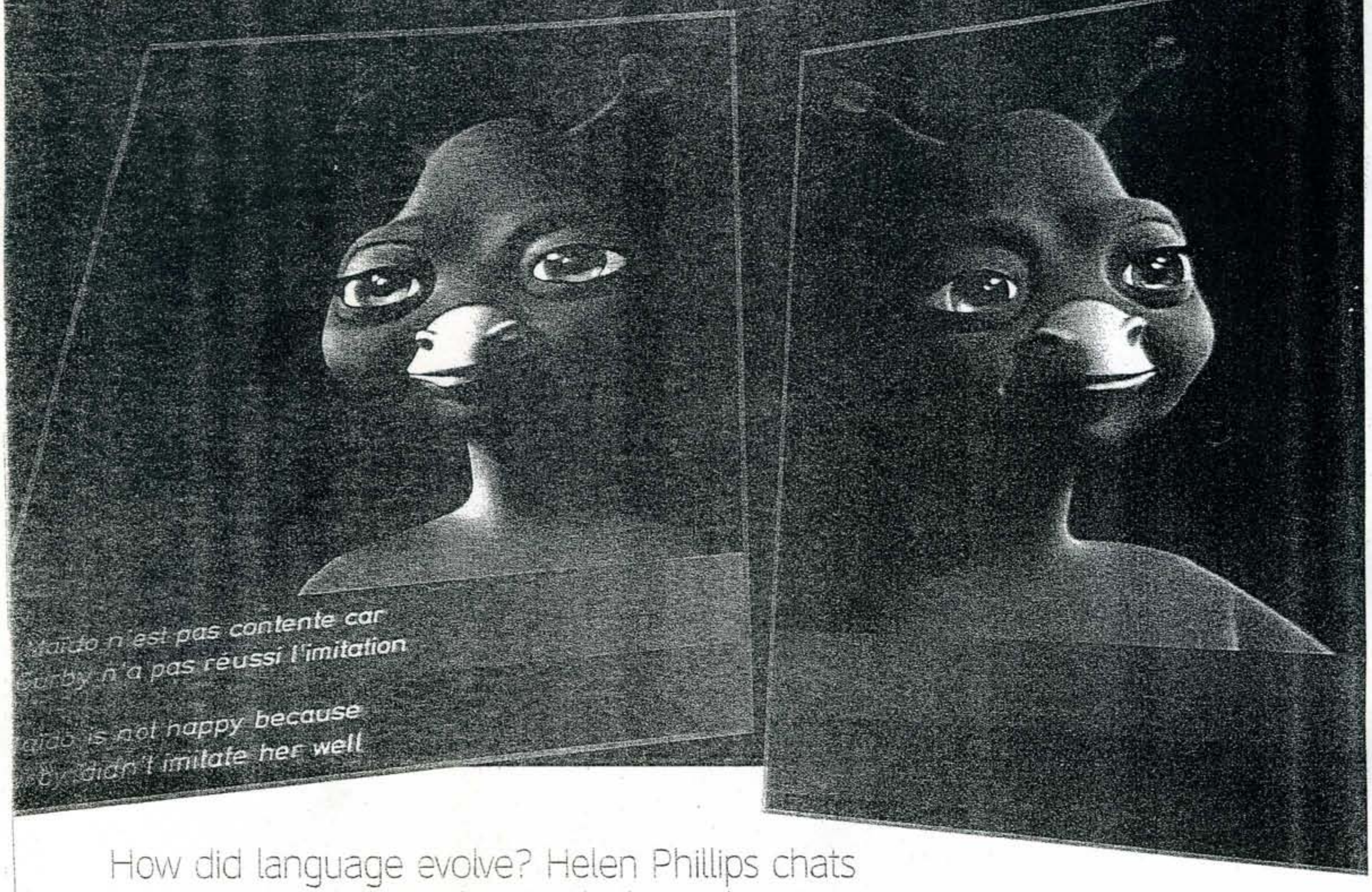
FIRST WORDS



How did language evolve? Helen Phillips chats to some cute robots that might have the answer

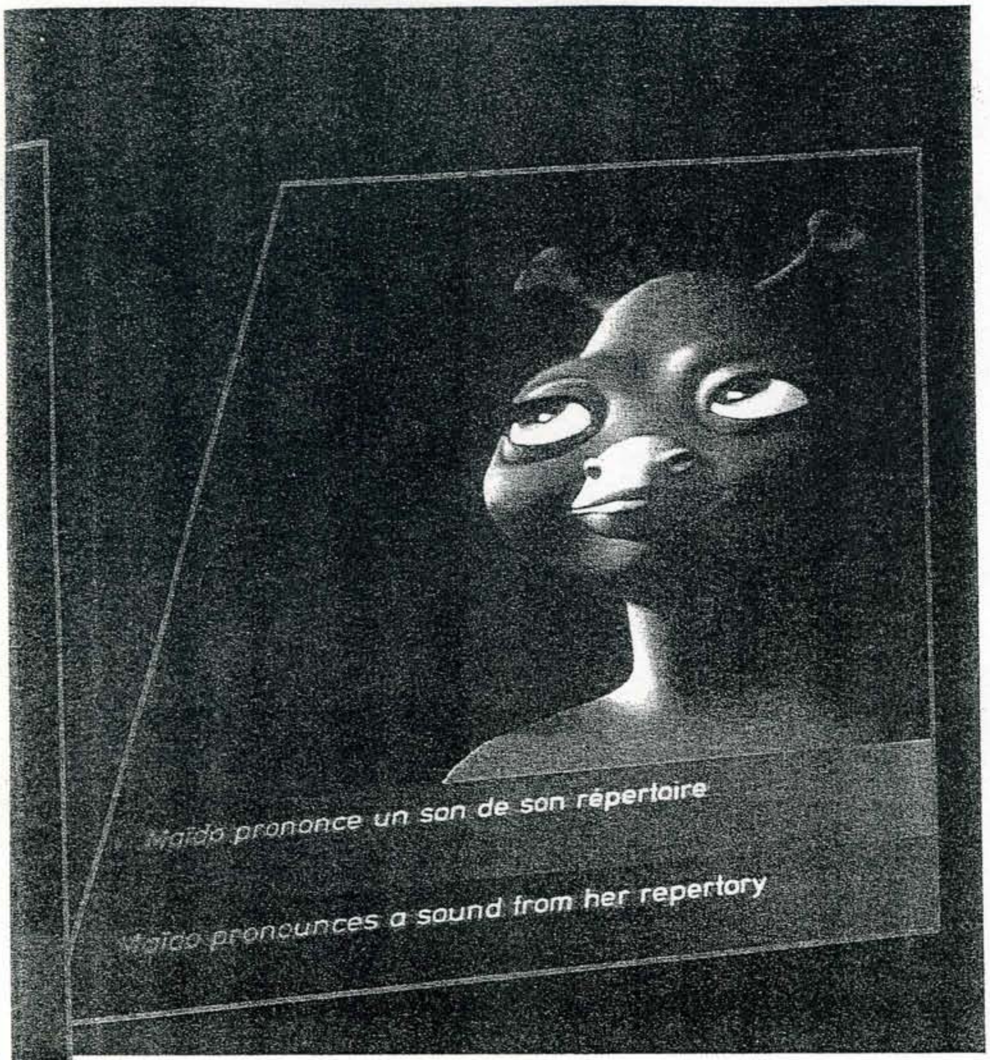
THE SONY Computer Science Laboratory in Paris is a cosmopolitan sort of place. Here an international team of researchers converse in English, French and Japanese. But the air is also full of more exotic voices babbling, or uttering strange words such as "wabaku" and half-recognisable phrases like "push red wa blue ko". These are Luc Steels's talking robots. Even the most accomplished linguist will have problems making polite conversation with them, because they don't speak any language we know. Instead they invent their own.

For decades, linguists, anthropologists and biologists have argued about what allowed our ancestors to evolve something as complex and elegant as language. What sort of brain would they have needed? Followers of linguist Noam Chomsky believe that some kind of linguistic rules must be encoded in our genes and brains. So to get language off the ground in the first place, specialised linguistic structures must have evolved in our brains. His opponents argue that picking up language is simply a matter of learning, and that given enough examples we can extract meaning, rules and order from what we hear, through a sort of subconscious statistical analysis. So language evolution is more about developing the right learning and rule-extracting skills (*New Scientist*, 21 August 1999, p 36). But years spent listening to electronic voices has convinced Steels and his colleagues that there is a third way.

Many times over they have heard new languages evolve in computers that are not programmed either with the equivalent of an innate linguistic sense or statistical powers. Instead, for each new language, rules are gradually invented, negotiated, built upon and spread by pairs of robots talking to and learning from one another. "Language is a complex adaptive system," says Steels. "It's like a living thing." It self-organises and spreads like a virus.

Over the past five years Steels and his colleagues, both in Paris and at the Free University in Brussels where he's head of the artificial intelligence lab, have witnessed the emergence of language-like communication between robots. His machines have created basic speech sounds and put these together to form words that have meaning. Other researchers are impressed by the flashy presentation and the neat technology, but few have been persuaded that Steels has got anywhere near the complexity of human language. But that view looks set to change after Steels's presentation at this week's Evolution of Language conference at Harvard University. There he revealed that his robots have done what some people thought impossible: they have evolved a form of grammar. "It's a bit of a bombshell," says

Pascaud/Eurelios



Chris Knight, an anthropologist at the University of East London, who is one of the conference organisers. "With minimal programming, just goals, agendas and the desire to form relationships, you end with something a little bit like language."

Steels's work stems from an idea he had back in 1995. His starting point is that language, particularly in its earliest phases, develops out of shared experiences, and so has to be grounded in the real world. In other words, there has to be something to talk about. So he began designing ways for robots with "bodies" and "senses" to have joint experiences, and programmed them with the desire to communicate these. "He has done something no one else has done," says James Hurford, head of the Language Evolution and Computation Research Unit at the University of Edinburgh and another conference organiser. "He's actually got real robots to have some communicative interaction about real things." And as the robots have been finding ways to speak to one another, Steels has been able to investigate what kind of brains, memories and other abilities are needed for language to evolve.

Most of the experiments so far have focused on developing words to express objects the robots can see. This is harder than it sounds, because the robots not only have to invent new words and store the meanings, they also have to work out concepts upon which to build the words in the first place. And they have very little to work with. A camera allows them to see the world around them—which consists of a white board covered with colourful geometric shapes—and they can point. Their "brains" include a standard text-to-speech module, some memory and image-analysis software that can categorise what the camera "sees" in terms of colour, shape, size, texture, position and other sensory dimensions.

Steels programs pairs of robots to play a cooperative guessing game. One takes the role of "speaker", chooses an object it can see and then tries to express which one it is "thinking" about. It can use any concept, but it has to devise ways to convert the information from the sensory channels into useful categories such as "left" or "red". Sometimes a combination is needed. If there's a red square and a red circle, "the red one" obviously won't be enough. If the robot has no word already in its memory—and they don't have any to begin with—it will invent something by combining strings of sounds from a basic repertoire of syllables.

The other robot, the "listener", must then try to figure out what the speaker is talking about. It signals by pointing, and will guess if it has no memory of the word it hears. If it's right, the speaker confirms with an "OK", and the word and meaning are reinforced in both their memories (see Diagram, p 26). In this way each robot develops a lexicon—a tree-like mental look-up table of words and meanings. New words and concepts grow like new branches on the tree, and those that are wrong or not used wither away. In one of the largest experiments, a population of robots managed to develop around 8000 words (see "Vox pop", p 27).

Clearly it's not exactly how humans began using words. For a start, the robots' speech sounds are all given, while our ancestors must have evolved the ability to produce different sounds. And the robots are programmed to interact in a far more regimented way than any ancient humans. But, says Steels, it does say important things about how language can evolve. For a start, the robots don't develop concepts, rules and hypotheses statistically from vast numbers of examples. They can make an educated guess. We also have a rich capacity to guess what other people might mean, says Steels. "In my view this is the essence of language learning and an absolute prerequisite for the origins of language."

But do the robots need the sorts of specialised innate language skills that Chomsky spoke of? When it comes down to it, says Hurford, the programs that drive the robots and play the games are quite specific. "You could say that corresponds to something that you've built in, so in some sense it's innate," he says. Steels agrees, although he dislikes calling his programs an innate predisposition for language. He accepts that robots have abilities that could be called innate, such as a visual perception system that allows them to pick out one object from another and recognise shapes. They also have the desire to interact and be understood, and the ability to engage the attention of others. "There are obviously a lot of things, but they are very general," he says. Our ancestors would not have had to evolve these abilities specifically for language.

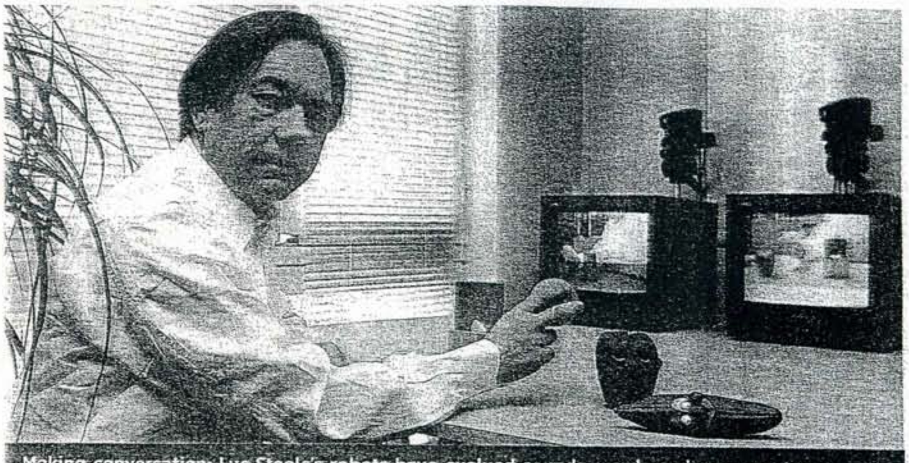
Perhaps the most radical claim Steels is making on the basis of these experiments is that the robots are developing words and concepts together. This sets his ideas apart from those who argue that statistical learning is all-important. While these researchers believe there's enough statistical regularity in the world to extract a set of basic "natural" concepts, the robots clearly do something different. They invent their own categories, which can only come from the sensory information they gain. By experience and guessing, those concepts develop word labels, such as "wabaku" and are passed on to other agents. "My idea is that language and meaning co-evolve," says Steels. "You cannot just have a concept and then develop a language label—the two are intimately linked."

This is a fascinating assertion, says Hurford. "The robots do in some sense develop their own conceptual framework. I'm sure he's right." It's a totally different idea to a lot of traditional theory. As far back as Plato there have been philosophers who argued that concepts are simply waiting to be discovered. And many linguists believe that language cannot emerge until we have an understanding of these universal concepts. Steels's picture is different. "There are many ways to view the world," he says. "But I think that language helps society see things in a similar way. It's a coordinating force." Knight agrees. "Language is clearly social, there's no question about it," he says. "Signals, sounds and concepts are socially generated."

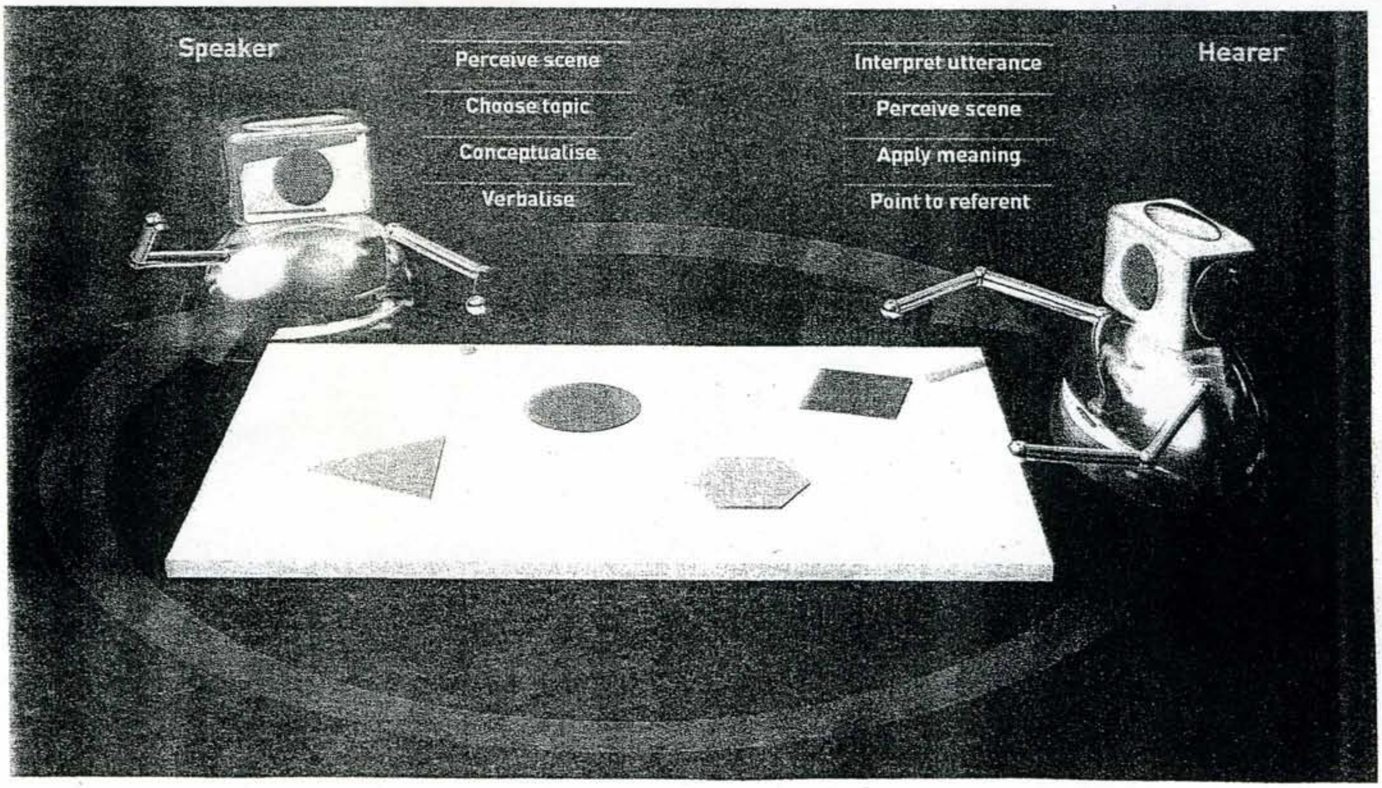
Steels points to the way we name colours as a good example. Because they fall on a spectrum there is no reason why we should split them in any particular way. Our visual system constrains us to some extent, but it doesn't explain entirely the categories we use. Culturally we came to a consensus, says Steels. He points to studies by Jules Davidoff of Goldsmiths College, University of London, and his colleagues, who describe how different cultures use entirely different colour categories. For example, English speakers recognise eight—red, orange, yellow, pink, green, blue, purple and brown. But in the language of the Dani people of Irian Jaya in Indonesia there are only two words used to express colours. And the Berinmo of Papua New Guinea have five categories (*Nature*, vol 398, p 203).

How words acquire meaning is fascinating stuff, but it's only a small part of language. It's grammar that has caused the most controversy. While there's general agreement that words and sounds are simple enough to be copied and memorised, many linguists, not just those who side with Chomsky, believe that grammar is just too complex and messy and that the basic rules must be innate. Steels disagrees. His latest experiments show that by watching and describing things happening, and so guessing what the descriptions mean, his agents can evolve at least one form of grammar, without any pre-set rules.

The game is similar to the word-creation experiment, but this time the robots watch dynamic situations, such as a hand grasping a ball. This has allowed Steels to explore the emergence of case grammar, which is what we use in describing the relationship



Making conversation: Luc Steels's robots have evolved sounds, words and even grammar



between objects. If you just have two objects, "ball" and "hand", with an action "grasp", without rules, you quickly run into problems in communicating unambiguously. Case grammar in English is largely done with word order—so "dog bites man" has a totally different meaning from "man bites dog". Other languages use word endings or additional words, known as case markers.

In trying to solve such ambiguities, Steels's robots have the ability to invent their own rules. "Push red wa blue ko" was a construction used to communicate that "someone pushes a red object against a blue object". By seeing the same action as the speaker, the listener can work out what role the unknown words "wa" and "ko" are playing. Because the robots are programmed to be economical with their memories, eventually they develop general rules for using these sorts of case markers, that might, for example, label who is "doing", and who is "being done to".

Rewriting the rules

"It's the first time with computer programs and real computer vision we've seen a real grammatical system emerging," says Steels. More importantly it shows that the basic categories usually associated with case grammar are not innate, he adds. Now he plans to test whether his talking heads can evolve the use of different tenses—for which they'll have to develop concepts about the temporal relationships between things they see.

This work has made Steels think about grammar in a different way to many linguists. "It's no frustrating set of rules to be used rigidly and to be grammatically correct. It's purely to help us to understand," he says. This has convinced him that language is constantly evolving—not just the words we use but also, over time, the whole rule basis upon which our unique form of communication is constructed. "Chomsky has a static view of language," says Steels. Chomsky insists that some fundamental rules of grammar are common to all languages, because they are in some sense hard-wired into our brains. Steels just doesn't buy that.

But if he's hoping to convince supporters of an innate language ability to think again, he's in for a tough time. "Given what we know about real human languages, and the way real human children acquire them, only hypotheses that attribute some innate specialisation for language to children can account for the data," says Steven Pinker from MIT. He also points out that genes linked with language are now being discovered. One study, for example, showed that identical twins have greater similarities in some of their grammatical habits than ordinary siblings or fraternal twins. Another revealed a gene linked to a specific language defect (*Nature*, vol 413, p 465). "It is incoherent to say that there are no genetic factors, because evolution, by definition, is a change in genetic composition over time," says Pinker.

And Jeff Elman of the University of California at San Diego points out that Steels's work might not rule out the statistical learning ideas at all. He might be exactly right about how language evolved in the first place, says Elman. But that doesn't mean that children don't then learn language by induction and example. The two ideas are not mutually exclusive.

But so far, the grammar skills of Steels's robots are not sufficiently human-like and the conditions not realistic enough to say he's found the roots of language. And though he acknowledges that simulations like his will never be able to prove how language came about, only to test under what conditions it can, Steels is already addressing the problems of realism. In the first experiments his robots were programmed with basic speech sounds, but now they can evolve their own. The researchers have found that if they give the computers a realistic virtual human vocal tract, and a virtual ear, they develop sounds remarkably like the ones we use.

And instead of playing games where the interaction follows a strict pattern, Pierre-Yves Oudeyer, a colleague of Steels, has developed a system of emotions, to give the robots a desire to seek out interactions and a sense of "fulfilment" after communicating successfully. It's much more realistic than the ritualised games. "Here they are truly autonomous," says Oudeyer. The emotion-led games are interactions rather like mother and baby, learning how to engage each other's attention, create sounds, imitate and reinforce successful copying. It makes them chatter and babble almost continuously.

But perhaps the only real proof will come when all the pieces of the language-evolution puzzle have been put together, to see whether it's possible to simulate the whole process, progressing through a stage rather like babbling, on to simple language, then a lexicon, and finally grammar. Even then, we won't know if that's what really happened in our own evolutionary history—only that it might have.

One thing's certain. Robots and computers can talk to each other, whether we like it or not. How long before they are teaching us to speak a new language? It's an interesting by-product of this academic curiosity. No wonder Steels says: "Linguists find this work all very odd."

More information can be found at www.csl.sony.fr or http://arti.vub.ac.be

Vox pop

TO TEST how language develops and spreads in a population, Luc Steels from the Sony Research Laboratories in Paris needed to find a way to do a large-scale social experiment—with robots. He wanted the robots to be able to meet and travel, while still allowing the researchers to control their surroundings. Steels, along with Angus McIntyre and Frederic Kaplan, hit on the idea of teleportation.

They realised that they could build just a few robot bodies but many robot brains. The brains were in the form of software "agents" that could be teleported, or downloaded into a body over the Internet. In this way hundreds, even thousands, of different robotic agents could meet and interact in pairs in a few controlled locations all over the world.

The robotic hardware—a camera to see the surroundings, and the ability to point, speak and hear—was stationed in various public places in Brussels, Paris, London, Tokyo and elsewhere. At each location a series of agents would use the bodies to play a language game in which they had to devise words to describe geometric shapes on a white background. Agents would then teleport off to another location with a different array of shapes, so that in a fraction of a second words learned in Tokyo might be used in Paris.

The experiment became a public one. People were invited to create an agent through a website and choose where it went to play its games. They could follow its progress and even act as a speaker or listener themselves, teaching agents words, or trying to learn from them.

Soon these globetrotting robots were using English, German, French and Japanese words, as well as ones they had invented themselves. Briefly, some of the more isolated groups would develop different dialects and languages, but with more interaction, words and meanings mingled and spread. At its height, the population was about 3000 strong, using maybe 8000 words, 300 of which were stable and almost universally understood.

Then the inevitable happened. Hackers began to teach agents to swear and the foul language spread like a virus through the population. You can't have robots swearing in public, so the experiment had to be stopped. Steels admits they may have been naive. But, he adds, it showed us one thing: the more robots interact with people, the more they talk like us.