

How Your Genome Helps You Speak

INAUGURAL SPEECH BY PROF. DR. SIMON E. FISHER

*in au
gurele
redo*

change perspective

Radboud Universiteit



INAUGURAL SPEECH

PROF. DR. SIMON E. FISHER



Simon E. Fisher is director of the Max Planck Institute for Psycholinguistics and Professor of Language and Genetics at the Donders Institute for Brain, Cognition and Behaviour in Nijmegen, the Netherlands. Simon obtained his Natural Sciences degree at Trinity Hall, Cambridge University,

followed by a DPhil at the Genetics Unit of the Biochemistry Department, Oxford University. For his postdoctoral research he joined Prof. Anthony Monaco's group at the Wellcome Trust Centre for Human Genetics (WTCHG) in Oxford, and worked on identifying genetic factors that contribute to dyslexia and speech and language disorders. He and his colleagues discovered *FOXP2*, the first case of a gene mutated in speech and language impairment. In 2002, Simon was awarded with a Royal Society Research Fellowship and became head of his own laboratory at the WTCHG, where he used state-of-the-art methods to uncover how language-related genes influence the brain. From 2007-10 Simon was also the Isobel Laing Fellow in Biomedical Sciences at Oriel College, Oxford. In 2010 he became director of a new department devoted to "Language and Genetics" at the Max Planck Institute. Simon is an elected fellow of the Society of Biology and has been awarded the Francis Crick Prize Lecture, and the Eric Kandel Young Neuroscientists Prize.

Radboud Universiteit



HOW YOUR GENOME HELPS YOU SPEAK

How Your Genome Helps You Speak

Inaugural speech delivered at the acceptance of the post of Professor by special appointment of Language and Genetics at the Radboud University Faculty of Science, on Thursday 26 September 2013

by Prof. Dr. Simon E. Fisher

Photography cover: Bert Beelen
Production: Radboud University, Facility Services

ISBN/EAN number: 978-90-9028871-0

© Prof. Dr. Simon E. Fisher, Nijmegen, 2013

No part of this publication may be reproduced and/or published in print, photocopy, microfilm audio tape or in any other manner without prior written permission of the copyright holder.

*Mijnheer de rector magnificus, dames en heren,
Zeer gewaardeerde toehoorders,*

Today, I want to share with you some intriguing research findings, ones that have played a very significant role in shaping my thinking about language and about the world. The results come from an exciting project that I started towards the end of the 1990s. I say 'I', but as with most scientific endeavours, this work could not have been done alone. The results that I'm going to report to you now involved a particularly close collaboration with one colleague, my wife Vicky.

This was a longitudinal, developmental study with two child participants, following them from birth. I'm going to refer to them as TJF and SAF to preserve their anonymity. Following the birth of each participant, my colleague and I noticed some quite astonishing and puzzling developments, which were strikingly consistent across the different participants. To quote from our research notes:

'During the first few months of life, each of the subjects made unusual noises with their vocal apparatus, which became more and more complex and, after a while, we began to suspect that these odd vocalisations might correspond to words. As time went on, the participants' understanding and expression of words became increasingly sophisticated. Not only did their vocabularies expand at an alarming rate, but they began to recombine different words to form novel utterances, using these sentences to convey a host of different ideas, not only relating to the present, but also the past, the future and even abstract concepts. This would have been remarkable enough, but we were struck by the fact that we had not at any point provided any formal tuition in this area. It was almost as though the subjects of our study had soaked up language from the environment around them.'

Now I have to admit, it's possible that Vicky and I were not the first to have noticed this strange phenomenon in our children. In fact, this type of natural 'research study' has been carried out billions and billions of times across the world, over and over again, with similar outcomes.

There is little doubt that there is something mysterious going on. As with other areas of biology, Charles Darwin was able to capture this mystery very elegantly in his prose. He said that: *'Language is an art, like brewing or baking. It certainly is not a true instinct, for every language has to be learnt. It differs, however, widely from all ordinary arts, for man has an instinctive tendency to speak, as we see in the babble of our young children: whilst no child has an instinctive tendency to brew, bake, or write.'* (Darwin, 1871; p.55)

Darwin refers to spoken language as an art. I think it's more like an amazing magic trick. Each of us is able to take the thoughts that are sitting in our own heads and pass them over into the head of another person. The way we do this is remarkable. Think about what's happening now: I'm coordinating the rapid movements of my larynx, tongue, jaw and mouth. In doing so, I'm pushing the air in front of my face backwards and forwards in such a way that I'm implanting in your brains these ideas about how wonderful language is.

But that's not all. Let's return to Darwin for a minute, and his vision of the natural world as a huge tree of life with many, many branches. How could it be that, in this whole tree of life, there is just a single twig with the capacity for spoken language? Why should this be confined to just one species – our own? As it happens, there are other species that learn their vocalisations by listening to the vocalisations of conspecifics. We are not alone in that respect. Suspected vocal learners include seals, dolphins, whales, bats and elephants, and a number of bird species, including hummingbirds and zebra finches (Jarvis, 2004). Take the zebra finch as an example. After male birds of this species hatch, they go through a stage that is very much like babbling, and then they learn to sing a specific song that they hear from an adult tutor bird. Vocal learning is not the same as language, but it is a key capacity that we depend on when we are acquiring speech. It is curious to see that the vocal-learning species – bats, some birds, certain mammals and so on – are found scattered through the tree of life, in each case closely surrounded by non-vocal learners. This suggests that vocal learning is a skill that has evolved separately a number of different times in evolutionary history.

Let's zoom in on our closest living relatives, chimpanzees. It is well established that chimpanzees lack capacities for vocal learning. In fact, they are very bad at vocal control in general. But, of course, speech is not the only way that we can use language. What about non-spoken forms, such as sign language? Interestingly, studies of sign language provide strong further evidence for language being deeply rooted in biology (Petitto & Marentette, 1991). Deaf babies will babble using their hands. Groups of deaf children with limited access to sign language have been shown to develop their own language-like gesture systems with many of the hallmarks of spoken language (Senghas et al., 2004). So, could a chimpanzee, even though it lacks control of vocalisations, still become proficient at language use, but through sign instead of speech? Unfortunately, the answer to this question is: no. Even with intensive tuition from birth, a chimpanzee can't remotely match the capacity of a young human infant. A good illustration of this comes from studies of the mischievously named chimpanzee Nim Chimpsky (Terrace et al., 1979). It's open to debate, but Nim was able to learn maybe 100 or so different signs, which is not a vast number compared to the typical vocabulary of an infant.

What's more, Nim always used these to elicit a desired outcome, while a human child goes far beyond this, becoming adept at using language to generate and express thoughts and ideas. Typical examples of Nim's longer sentences are 'grape eat, Nim eat', 'tickle me, Nim play' and so on. As far as I'm aware, no chimpanzee has ever asked something like 'why is the sky blue?' or 'why am I here?', and these are questions that children ask all the time.

So, the origins and basis of human language are something of an enigma. Perhaps the answers might be buried somewhere in our genetic makeup. Each of us has 22 chromosome pairs, one chromosome coming from our mother, and one chromosome coming from our father. In the case of a male, we also have an X-chromosome and a Y-chromosome, while the female has two X-chromosomes. These chromosomes are all contained in the nucleus of almost every cell in your body, and they make up your genome. Effectively each chromosome is one extremely long string of letters of DNA, bound up with other material. DNA makes use of a four-letter code, with A's, T's, G's and C's, strung together. If you add up all the different letters across your genome, this is about 3 billion letters long. Of course, it's obvious that the particular language that you learn could not be encoded in your DNA. A child who grows up in Japan, surrounded by Japanese speakers, becomes fluent in Japanese, but the same child growing up surrounded by Dutch speakers would become fluent in Dutch. And if you aren't exposed to a language at all, you won't acquire it. But at a deeper level, genes lie at the very heart of this process, helping to build a brain that is finely tuned to soak up speech and language skills from the social environment (Fisher, 2013).

To understand why this is possible, we need to think a bit about the way that information in DNA is used to direct biological processes. For many stretches of DNA in our genomes, we still don't understand what they do. The bits that we know the most about are the genes that code for proteins. We carry maybe 20,000 or so of these different genes, and the way these work is very well established. A protein-coding gene consists of a defined stretch of DNA, containing a linear code of letters, which can get copied (transcribed) into another molecule, called messenger RNA. That messenger RNA molecule is basically just a copy of one of the strands of DNA. The only difference is that instead of letter T's there are letter U's, but, other than that, it's the same code. The machinery of the cell is able to use this messenger RNA molecule to read off the instructions for building a protein. There are 20 different types of amino acids that can get strung together in potentially infinite numbers of combinations. The sequence in the messenger RNA is used to designate which amino acid should be attached to which other amino acid, using a triplet code. For example, AUG is the codon that specifies insertion of a methionine residue, AGU specifies a serine residue and so on. The biological machinery in the cell knows what to do, and it strings all the appropriate amino acids together according to the information in the messenger RNA sequence.

When a string of amino acids comes together in this way, it folds up into a characteristic three-dimensional shape. The shape of the protein is determined by the sequence of amino acids, and the function of the protein is determined by its shape. This is an amazing system when you think about it; the genome is able to use a simple linear code to describe highly complicated three-dimensional structures. The proteins that are built have many diverse jobs that they can do in the cells in your body. There are enzymes, receptors, signals, transporters, structural proteins and so on. Together they form the molecular machinery that makes you a human.

The other important thing to recognise is that each of us carries variations in our DNA sequences, individual points in particular genes where letters are different. Many are common changes that might be shared by a significant proportion of the population, while others are rare variants. Some might perhaps even be unique to you and your family. Sometimes these changes are silent; they don't affect how a protein will work. But certain types of changes might alter the shape of a protein, or the way that it's being switched on or off, or how much of it is made. This can lead to effects on your physiology, anatomy or behaviour. Such effects are extremely well recorded for many aspects of biology, so why not for human speech and language?

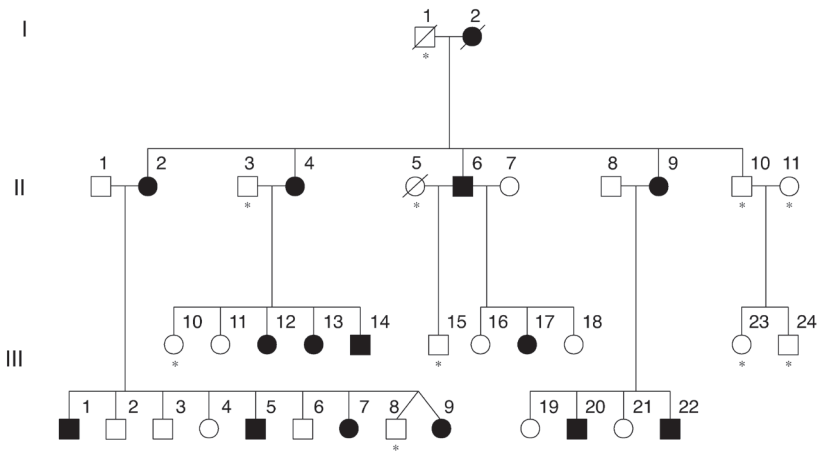


Figure 1: The KE family. About half the members of this three generation (I-III) family were found to be affected by a severe speech and language disorder. The pattern of inheritance suggested the disorder in this family may be caused by mutation of a single gene, motivating a genome-wide search to pinpoint the locus responsible. The affected individuals are indicated by filled symbols. Squares represent males and circles represent females. A line through a symbol indicates that individual was deceased at the time when the genetic study was carried out. Asterisks indicate individuals for whom DNA samples were not available.

To illustrate the first lead we had into this interesting question, I want to introduce you to an unusual family, known as the KE family (Figure 1), originally identified by our collaborators at the Institute of Child Health in London. Across the three generations shown in the family tree, there are a large number of people suffering from speech and language problems. In each generation, about half the children are affected, while half are unaffected. Based on the pattern of inheritance of the disorder, it looks like this might be down to just a single gene being damaged, which is very odd for something as complicated as speech and language. The most prominent problems that these children face are problems with speech coordination, a disorder known as 'childhood apraxia of speech' or 'developmental verbal dyspraxia'. They find it hard to make the sequences of movements, with mouth, tongue, lips and so on, that are important for producing fluent speech (Watkins et al., 2002a). The difficulties persist into adulthood, and you can see this clearly if you give the adults in this family the task of repeating a word like 'catastrophe' a few times. The affected people do really badly with this, and they make different types of errors each time they repeat the word. They can produce the individual sounds, but they have problems selecting the right sequence to make. On top of this, they have wide-ranging impairments with multiple aspects of language, not just affecting expression but also extending to the understanding of language, even affecting grammatical processing.

What was particularly exciting about the KE family disorder was the fact that it could potentially be due to just a single gene being damaged, and nothing like this had been seen before. I was working in Tony Monaco's research group in Oxford, studying the DNA from this family, and we were able to zoom in on a region on chromosome 7 which we suspected contained the gene that was responsible (Fisher et al., 1998). Eventually, we pinpointed a single tiny change in their DNA, which we believed was causing the problem. This was work that I carried out together with Cecilia Lai, a PhD student in Tony's laboratory. We discovered that all of the affected people in the three generations of the KE family carry a change: a mutation of just one single letter at one point in this gene, from a G to an A. This was a tiny change at the level of DNA, but one with big consequences for the speech and language development of the people who carry it (Lai et al., 2001).

As mentioned earlier, many genes code for proteins. We could quickly determine what type of protein was being made by this new gene that we had identified by reading off its DNA sequence and figuring out the amino-acid sequence of the encoded protein. We discovered that the protein contained a special section, called a 'FOX' domain. Proteins include different parts that form different shapes, and we could say something about the shape that this domain would form, because it had been found in other types of proteins. It forms something called a 'winged helix', where it has three helical

domains and loops between them. When we looked further, we saw that, in the KE family, the little mutation we found, a change to a single DNA letter, led to a change in the sequence of the encoded protein, right in the middle of one of the helices of this FOX domain (Lai et al., 2001). I will return to this later to explain more about why that is important. We were able to show that changing the structure of the protein in this way was enough to interfere with how the protein worked. And we hypothesised that this was enough to cause the disorder that we were seeing in the family.

We have gone on to identify other families, nowhere near as large as the KE family, who have similar problems and again have mutations in this particular gene (a gene given the name *FOXP2*). It's worth pointing out that the specific mutation of the KE family has never been seen in another family, but other damaging mutations elsewhere in the gene have been found. For example, in 2005, we discovered a separate mutation in a relatively small family where there were three people affected (compared to the 15 people affected in the KE family). In fact, this smaller family have a very interesting type of mutation which puts a stop signal right in the middle of the protein (MacDermot et al., 2005). Instead of making a full-length *FOXP2* protein, the affected people in this family make a truncated one that doesn't work properly, because it is missing important parts (such as the FOX domain itself). In the years that followed, a handful of other *FOXP2* mutations have been found in rare cases of speech and language disorder, some of which have yet to be published.

There are different types of changes that seem to implicate *FOXP2* in different families. I've already discussed how alteration of just a single DNA letter can lead to a big disruption of the encoded protein. Sometimes, rather than there being a tiny mutation of one letter, a whole chromosome is damaged in some way. For instance, a person might carry a translocation, where one section of one chromosome has broken off and become attached to another section of another chromosome, so that the two chromosomes have exchanged material. In fact, a person carrying this kind of change may carry all the genetic material that everybody else has, but they have a break in their chromosomes. If that break lies in the middle of a gene, it can damage the way that the gene works. In another type of change, known as a deletion, somebody could miss a whole gene because an entire section of DNA has been lost.

We now know of several cases where translocations or deletions impact on *FOXP2* and cause speech and language disorders, with similar cognitive profiles to that which we saw originally in the KE family. The bottom line is that two functioning copies of the *FOXP2* gene are necessary to acquire proficient spoken language (Fisher & Scharff, 2009). Damage to one copy is sufficient to cause a severe speech and language disorder.

Nevertheless, as I've shown you, there are only a small handful of these *FOXP2* mutations in the world. So *FOXP2* mutations can't be the whole story. What about other genes involved in speech and language? How can we find these? As mentioned earlier, each of us is carrying around a great many different types of DNA variants. Some of these are relatively common in the population and might put us at increased risk of different types of disorders. Importantly, with current technologies, we are able to scan all these known common variants because they are already documented. Using a DNA chip, we can look at hundreds of thousands of common variants on each chromosome. We can ask whether there are particular variants that are found more often in children with language disorders than in unaffected people. We use rigorous statistical approaches to determine that the association is real and not simply a chance finding. We can also target individual genes or variants in them that are already known to be interesting. For example, we might examine common variants in the *FOXP2* gene, to see whether they are important for common forms of language disorders. In the Language and Genetics department, we have a number of people doing this kind of work on common variants, including Clyde Francks, Tulio Guadalupe, Alessandro Gialluisi and Amaia Carrion-Castillo. To carry out our gene-mapping work on language impairment and dyslexia, we collaborate with quite a few different groups in the UK and America and also here in the Netherlands. In particular, we have started a network of Dutch researchers who focus on developmental language disorders. I am also working together with researchers of the Dutch Dyslexia Programme (DDP), a project studying cases of dyslexia in the Dutch population, which has been running for over 10 years.

So, overall, we're doing a lot of work in various different cohorts looking at contributions of common gene variants. What about the impact of rare unknown mutations? A problem with screening people (genotyping them) using DNA chips is that we need to already know the different gene variants that we want to examine. In the case of *FOXP2*, the specific mutation we discovered in the KE family is present only in that one family. (As I've explained, other families have different rare mutations.) We might never have found the mutation in the KE family if we hadn't already had clues pointing to *FOXP2*. How can we find these kinds of rare effects that might just be popping up in one family? Is there some way to sift through all the genes, every single gene that a person carries?

I want to take a slight detour now to provide a bit of perspective. Back in 1990, when I was an undergraduate student, I read a paper by Jim Watson on 'The human genome project: past, present and future' (Watson, 1990). Scientists had just begun an endeavour to sequence a whole human genome, an incredible thing to be able to do. It was compared at the time to President John F. Kennedy planning to send a man to the

moon, so this was a big deal, and I was very excited when I first read the article. What happened next? More than 10 years of hard work followed, involving thousands of machines and researchers (I was lucky enough to be one of them), trying to assemble just a single human genome at a cost of at least \$3 billion. After all this time, effort and money, in the year 2000, Craig Venter, Francis Collins, John Sulston and colleagues announced the completion of a first draft sequence of the human genome. Let's fast forward another 10 years after this milestone. Still, in 2010, there are many lingering mysteries about how the information of the genome can translate into a thinking, feeling, remembering, talking person. So maybe the advances in our understanding as a consequence of the human genome project have not been so dramatic? On the other hand, by the year 2010, some other things have happened that are having an enormous influence, in particular related to the challenge of sequencing more human genomes. At this point, one or two researchers can put in a few days of work, using just a single sequencing machine, spend around \$3000 and thereby sequence all of the protein-coding part of the genome (1–2% of it, a portion known as the 'exome'). As scientists started using these new sequencing technologies, they began uncovering all sorts of interesting things. In fact, some of the people at the forefront of this extraordinary new wave of discoveries were here in Nijmegen. Alex Hoischen, Han Brunner and Joris Veltman in the Radboud University Human Genetics department showed that you could use this exome sequencing to track down the causes of rare diseases that nobody had ever been able to figure out before (Hoischen et al., 2010). They identified unique gene variants in patients, mutations that were responsible for severe (often lethal) genetic disorders. We were very excited to be able to start collaborating with Alex, Han and Joris to apply these kinds of techniques to specific language impairment, in collaboration with Dianne Newbury from Oxford and Sylvia Chen and Clyde Francks from my department. Hopefully, we'll have some interesting findings to report from this work in the not too distant future.

We are also using these new sequencing technologies to look at multigenerational families who are a bit reminiscent of the KE family (that is, with many different relatives affected with a language-related disorder), with the idea that we might be able to find the responsible gene or genes quicker than in our earlier work. For example, Kate Kucera in our department is working in collaboration with David Skuse and Josie Briscoe, studying a family with an unusual kind of developmental deficit in linking semantic knowledge to language (Briscoe et al., 2012). Amaia Carrion-Castillo, together with Barbara Franke and Ben Maassen of the Dutch Dyslexia Programme, is applying the same kind of approach for a big family with dyslexia that has been previously studied in the Netherlands (de Kovel et al., 2004).

By now, nearing the end of 2013, it seems that we have already entered the era of personal genomics, as the costs of sequencing technologies have continued to plummet. It may be not so long in the future when you don't have to have an unusual disorder, or be famous and important, in order to have your whole genome sequenced. What does this mean? We will be uncovering more and more genetic information from people. We can expect to find many individual DNA variants that are important for all sorts of things, including, in our case, for studying speech and language.

One of the big problems is that there is a huge gap, actually an abyss, between the level of DNA and the level of what we're trying to understand about speech and language (Fisher, 2006). For the last part of my lecture, I want to give a brief sense of how we can fill in the different levels in between. We're going to consider the *FOXP2* story, because this is the case for which we have the most information, but the principles apply for any interesting gene we might find.

As I already said, the genes that code for proteins are the genes we like the most, because they are the easiest ones to study. In the case of *FOXP2*, we can ask: what does the resulting *FOXP2* protein actually do in the cell? I showed you earlier that the *FOXP2* protein has a special domain, the FOX domain, containing three helices. Remember that the affected members of the KE family have a mutation that makes one particular residue at this point of the protein, in the third helix, change from an arginine to histidine. It changes the shape of the helix and disturbs what it does. Normally, this domain sticks to DNA. That is, the *FOXP2* protein likes to go into the nucleus of the cell, find particular stretches of DNA and stick to them. Why would it want to do that? Well, although the genome contains 20,000 or so protein-coding genes, they all do different things and they can't all be switched on at the same time. It would be like an orchestra in which all the instruments were playing different tunes at the same time. In front of genes and within genes, they have regulatory sequences that help to keep them quiet, or get them to play when the time is right. From what we can tell, what *FOXP2* seems to do most of the time is to stick to regulatory sequences in other genes, acting to tune down the expression. We often say that it is like a genetic dimmer switch. This belongs to a very important class of proteins: factors that work to orchestrate the music of the cell. *FOXP2* is regulating how other genes are being switched on or off, which leads to an interesting research avenue that we have been able to pursue in the decades since we first discovered the gene.

What are the target genes that *FOXP2* regulates? Do they do interesting things in the cell? Are they linked to language impairments? Sonja Vernes has done lots of really beautiful work on these questions over the years, starting with a very nice paper in 2007, the first to identify *FOXP2* targets (Vernes et al., 2007). What is intriguing about

these targets, as we have started to look through them, is that some have roles in the central nervous system and seem to be linked to other disorders and other traits (Vernes et al., 2008). One example is a gene called *CNTNAP2*, (contactin-associated-protein-like-2) which seems to be important in the way that the early nervous system gets developed, as well as contributing to how neurons are able to communicate with each other in the adult. This gene has also been implicated in neurodevelopmental disorders such as autism, schizophrenia and epilepsy. Another target gene that has been studied is called *DISC1*, which stands for 'disrupted in schizophrenia', while a couple of other targets (*SRPX2* and *uPAR*) have been associated with epilepsy. So, some of the genes that *FOXP2* is regulating are themselves implicated in various neurodevelopmental traits.

We might also ask: what about other proteins that interact with *FOXP2*? This is a question that is being investigated by Pelagia Derizioti, Elliot Sollis, Sara Busquets Estruch and Sarah Graham in our department. We are working on several different proteins at the moment, some of which have been linked to neurodevelopmental disorders. I want to give you just a brief example of one of these, which again speaks to the importance of *FOXP2* in this network of genes and proteins. The example is *FOXP1*, a protein that sticks directly to *FOXP2* and is able to sit on DNA in much the same way as *FOXP2* does. Pelagia has shown that *FOXP1* (just like *FOXP2*) targets the *CNTNAP2* gene, to down-regulate its expression (O'Roak et al., 2011). This finding is interesting in the context of one of the big DNA-sequencing studies that was done recently, by Evan Eichler and Brian O'Roak at the University of Washington in Seattle. They identified a *FOXP1* disruption in a family in which one child is affected with autism and severe speech and language problems, while his siblings and parents are all unaffected (O'Roak et al., 2011). The affected child carries a mutation of *FOXP1*, which is a bit like some of the *FOXP2* mutations that we previously found, because it completely knocks out one copy of the gene. However, this child does not only have a mutation of *FOXP1*, he also carries a very rare mutation in the *CNTNAP2* gene, which we know is part of the same pathway that *FOXP1* and *FOXP2* are involved in. This *CNTNAP2* mutation is also found in his unaffected mum and in an unaffected sister. But the affected child has a double dose of damage – two parts of the same pathway are being hit at the same time. We think that is something we are going to see a lot of as we continue this kind of work.

The conclusion from all the protein work that we have been able to do is that *FOXP2* is a key regulatory factor in gene and protein networks that are involved in neurodevelopment. However, I haven't told you anything about what *FOXP2* could do to a neuron (a brain cell). We got some initial clues from studying the genes that *FOXP2* was regulating, because we found that many of these targets were involved in the ability of neurons to make neurites – processes that grow out of the cell body and

eventually develop into dendrites and axons, providing the connections between different neurons. Sonja Vernes went on to take cells from brains of mice that have no functional *Foxp2* (the mouse version of *FOXP2*), comparing them to cells that did have functional *Foxp2*. She looked at how neurites develop from these neurons, measuring things such as the numbers of processes that grow out from each neuron, the length of the processes and how much branching they have. Sonja found that, when there is no functional *Foxp2* at all, the neurons in which *Foxp2* is important make shorter processes with fewer branches (Vernes et al., 2011). Sonja now runs her own lab at the Max Planck Institute and she is doing lots of follow-up work on these findings, in collaboration with other colleagues in Nijmegen, looking closely at how *Foxp2* affects growth of axons, the main connections from one neuron to another. The findings already give us a hint that it is important for the way that neurons connect with each other during early brain development.

That brings us to the next level of explanation – neural circuitry. No neuron does a job completely by itself. There are lots of different types of neurons in the brain, and they are all wired up in very complicated ways. Can we study the level of neural circuitry and ask whether the *FOXP2* gene has any effect on the functions of circuits where it is usually expressed? To address this issue, we could again investigate mice carrying damage to the mouse *Foxp2* gene. This is work that was led by Matthias Groszer when he was in Oxford and that he is now following up in Paris. Matthias showed that, when mice have no functional *Foxp2* protein at all (that is, when both gene copies are damaged), it has really severe consequences (Groszer et al., 2008). In fact, they die a few weeks after birth. However, when mice have damage to just one *Foxp2* gene copy, they live long, healthy happy lives. You wouldn't notice anything strange about them, unless you start to look carefully at what they do when they are given running wheels to run on. For example, there is a type of tilted running wheel that we can put in the home cage of a mouse, enabling us to automatically record all the bouts of running and analyse the motor patterns. Mice with one damaged copy of *Foxp2* have normal baseline motor behaviour, but when they learn to run on these kinds of tilted running wheels (and on other similar systems, such as rotating rods that accelerate), they are much slower to learn. This is about making sequences of movements with paws and legs, and not about controlling the movements of orofacial muscles (as we do during speech). That may say something about the differences between what this gene is doing in a mouse and in a human. Still, speech is something that is highly dependent on learning to make rapid complicated motor sequences.

Catherine French and Rui Costa, collaborators in Portugal, have been able to delve even further into the relevant circuits in active living mice. While mice are learning to run on wheels or rotarods, Catherine and Rui can actually record from

individual neurons in parts of the brain where *Foxp2* is normally switched on (French et al., 2012). They find that some neurons selectively increase their activity when mice are running, while others become less active. This enables them to characterise such task-related neurons and ask what they are doing. In mice that have one damaged copy of *Foxp2*, the task-related neurons are generally overactive, even when the mice are not running. Moreover, although the number of task-related neurons in these mice is the same as normal littermates, many more of those neurons were down-regulated during running. By the way, these mice carry a mutation that is equivalent to that found in the KE family. We put the identical mutation of the KE family into a mouse and recorded from the neurons, in a way that is not possible in the human cases. These findings suggest that the gene is involved in plasticity in a mature animal (how these neural circuits are able to learn in adult life).

What do we know about effects of the *FOXP2* gene on human brains? Clues came initially from lovely work (work that I wasn't involved in) carried out by Kate Watkins in Faraneh Vargha-Khadem's group in London. This research was done before we had even found the gene. Kate and colleagues used magnetic resonance imaging to examine the grey matter density of different brain regions in the affected members of the KE family, comparing it to that of unaffected members (Watkins et al., 2002b). They identified places in the brain where there is significantly less or more grey matter in the affected KE members; we know now that this is a consequence of the *FOXP2* mutation that they carry. One of the more interesting sites is the caudate nucleus, part of the striatum in the basal ganglia. Other regions include the inferior frontal gyrus, overlapping Broca's area, which is a famous region that is important for language, as well as some motor areas. Later, after we discovered *FOXP2*, we found that the striatum is one of the places where this gene is switched on at the highest levels (Lai et al., 2003). It's an interesting convergence of brain-imaging and gene-expression studies. I've already shown you data from other studies pointing to effects of *FOXP2* on connections between different neurons. What do we know about the connections between the different regions in the KE family brains? There are methods such as diffusion-weighted imaging (another type of brain imaging) where you can actually trace the connections in people's brains. Kate Watkins, who is now in Oxford, is running work with the KE family to look for aberrant connections.

So those are studies searching for brain alterations in families such as the KE family, where people carry rare mutations. In Nijmegen, we also want to ask a slightly different question. Do changes in genes such as *FOXP2* have effects on the general population? We know that mutations of *FOXP2* that cause speech disorders are very rare. But lots of us carry different common variations of *FOXP2* - what do those do? This is the kind of approach where we look into genetic variability in the general

population in order to find out new things that will relate to disorders and mechanisms affecting brain development. In Nijmegen, a big team of us, led by Barbara Franke, are coordinating a programme called Cognomics (or cognition genomics). The idea is to have a large-scale programme where we can integrate genome information, functional/structural brain-imaging data and behavioural measures of memory, language and other aspects of cognition. Martine Hoogman, Tulio Guadalupe, Marcel Zwiers and Clyde Francks are now using this dataset, which is currently 1300 brains with matching genome data, to ask the question: Do common variants of *FOXP2* affect aspects of brain structure in the general population? And we have some hints about that that we'll be publishing soon.

Now, I want to highlight an interesting point about the work that I've been describing. For one section of the talk, I explained how we had to perform studies in mice. I told you that mice have their own very similar version of *FOXP2*, and outlined how we used mouse models to successfully study functions of this gene. What the research has shown is that *FOXP2* is not this magical thing that has popped up out of nowhere in humans to give us speech and language. Quite the opposite, *FOXP2* has been around for a long, long time; every vertebrate has a version of this gene. This is an old, old gene. It is doing interesting things in the brains of many different species (Fisher & Scharff, 2009). I just want to highlight a couple of examples. Constance Scharff in Berlin studies *FoxP2*, the bird version of the gene. *FoxP2* is found in many different bird species. Remember I said that some songbirds, such as zebra finches, learn their songs through a process of auditory-guided vocal learning. Constance knocked down the levels of *FoxP2* protein in a key part of the zebra finch brain and showed that it interferes with the zebra finch's vocal-learning abilities. This is extremely elegant work that Constance is still following up. We might even study the gene in flies. At first, this might seem crazy, but fruit flies have their own *FoxP* gene, and it appears to be important for brain development and function. The fruit fly is a very powerful model organism in genetics, in which you can study mechanisms and get valuable new insights. So we are in the process of working on this model system, together with Sonja Vernes and Annette Schenk here in Nijmegen.

These studies show how we share certain genetic mechanisms across distantly related species, from humans to mice to birds and even flies, pointing to deep evolutionary history. What about the issue of differences between humans and other species? Here we return to the question that I talked about right at the beginning. Why is it that human children are so good at communicating in all sorts of different ways, while chimpanzees seem unable to pick up sophisticated language, even under intensive tuition? Could *FOXP2* provide part of the explanation for this phenomenon? The people who answered this question were Svante Pääbo and Wolfi Enard from the Max

Planck Institute for Evolutionary Anthropology in Leipzig. Some years ago, they determined the sequence of *FOXP2* in different primate species and showed that chimpanzees have a slightly different version from humans (Enard et al., 2002). Two changes to *FOXP2*, altering the amino-acid sequence of the encoded protein, occurred on the human lineage after splitting from the chimpanzee. In subsequent work, Wolfi has used mouse-model systems to show that these changes impact on the way that *FOXP2* works in the brain (Enard et al., 2009). This is very interesting, because it suggests that the *FOXP2* changes might have contributed to cognitive differences between humans and chimpanzees. Nevertheless, it is certainly not the whole story, just one part of a bigger puzzle.

There is one especially fascinating message from all of this research. Our first glimpse into the genetic mechanisms that are involved in speech and language not only emphasises our differences from (for example) chimpanzees. It also illustrates our close connections to the rest of the animal kingdom, because the relevant genes have been around for a long time in evolutionary history. The findings suggest that our unique capacity for acquiring speech and language didn't pop up out of the blue, but was built on systems that are evolutionarily ancient (Fisher & Marcus, 2006).

In a similar, though much less dramatic way, my inauguration here at Radboud University has not appeared out of the blue, but is the result of a fairly long evolutionary process during my scientific career. So I want to end by acknowledging some of the factors along the way that have led to me standing in front of you today. Of course I would not be here at all, if it weren't for two people, my mother and father, Susan and Colin Fisher. I'm always telling students that the 'nature or nurture' debate is completely meaningless, because nature and nurture are tightly intertwined. So I want to thank my parents, not only for the genes, but also for allowing them to be expressed, and for nurturing their development over the past 40 or so years.

Talking of genes, my Oxford PhD supervisor, Ian Craig, introduced me to the joys of being an intrepid gene hunter. This is when I began to appreciate how dynamic and exciting scientific research really could be. I'm never going to forget the thrill of uncovering a new gene for the first time and reading off DNA sequences that no-one in the world had seen before. That was an amazing thing to experience. Tony Monaco, my supervisor and mentor during my post-doctoral years, has been a big inspiration. He taught me that the key thing is to concentrate on your science; do the best quality, the most rigorous science that you can, and keep your focus on that, and the rest just will fall into place. Tony is also a selfless promoter of junior scientists, and he was very generous in allowing me to build on the research that we had started in his lab, and to take it in new directions as head of my own group. From my experience as a post-doc

with Tony, I learned how essential it is to invest in the next generation of scientists. So I'm very grateful to the members of my Oxford lab, the post-docs, RAs, students and visitors over the eight years of its existence, who in their own ways continued to teach me how to be a mentor. I'm now in the rewarding position of seeing several of these people setting up their own labs and going in interesting new directions. It's a great thing to see, very exciting.

I would also like to thank the funding agencies that kept us going during the Oxford years. As well as the people from my own lab, the work I've described today has depended on invaluable contributions from many very talented and hard-working people. This has been a highly collaborative effort.

Now I want to go on to the people you can most 'blame' for my presence here in Nijmegen, my fellow directors from the MPI. Back in 2007, at a conference in Utrecht, a number of these fine individuals took me, along with a couple of the other speakers at that conference, to a secret venue, hidden away in the town, and we had a rather delicious, yet oddly secluded dinner, where they spoke of future plans for a completely new department with the challenge of linking genetics and language, and to be headed by some dashing young director, who was as yet unidentified. In all honesty, I really didn't imagine that, 3 years later, I would find myself sitting in an office in Nijmegen, beginning the process of establishing that very department. Life sometimes takes extraordinary twists and turns, but this has been a very welcome development. I can't think of a better place to be, for doing the science that I love the most. The MPI fosters innovative, imaginative and interdisciplinary research in a way that's rare to see elsewhere in the world. I want to thank my fellow directors, first for somehow being able to create my dream job, without me even asking them to, and second for being wonderful colleagues, really interesting, fun and enjoyable to work with. The research environment here in Nijmegen is absolutely superb; it benefits so much from these plentiful interactions between the MPI and the Radboud University, particularly in relation to the Donders Institute, and we've become much more than the sum of our parts. To my mind, the genetics and neuroscience studies that are being done here are among the best in the world, and I love being a part of it. I would like to thank my colleagues at the Donders for bringing me into the fold and for all the new opportunities for collaborative work that we are developing as a result. I'm grateful to Stan Gielen and Paul Tiesinga for being so welcoming as I've taken up my position as professor in the Faculty of Science. And to the members of the MPI Language & Genetics department (some of whom I have mentioned already), I want to thank all of you for what you've done so far, and for what you'll be doing in the future.

I want to end by returning to the very start of my lecture, and to the longitudinal study that I began with my absolute favourite collaborator of all, Vicky Fisher. You'll remember that this concerned our on-going investigations of two individuals, SAF and TJF. A couple of years ago, our investigations shifted focus away from first-language acquisition. We are now gathering lots and lots of data from these participants on the subject of second-language acquisition, in particular with respect to Dutch. So far, the results look very promising indeed. With that I will close. *Ik heb gezegd.*

BIBLIOGRAPHY

- Briscoe J, Chilvers R, Baldeweg T, Skuse D. 2012. A specific cognitive deficit within semantic cognition across a multi-generational family. *Proc Biol Sci*, 279:3652-61
- Darwin C. 1871. *The Descent of Man, and Selection in Relation to Sex*. New York: D. Appleton & Co
- de Kovel CG, Hol FA, Heister JG, Willemsen JJ, Sandkuijl LA, Franke B, Padberg GW. 2004. Genomewide scan identifies susceptibility locus for dyslexia on Xq27 in an extended Dutch family. *J Med Genet*, 41:652-7
- Enard W, Gehre S, Hammerschmidt K, Hölter SM, Blass T, Somel M, Brückner MK, Schreiweis C, Winter C, Sohr R, Becker L, Wiebe V, Nickel B, Giger T, Müller U, Groszer M, Adler T, Aguilar A, Bolle I, Calzada-Wack J, Dalke C, Ehrhardt N, Favor J, Fuchs H, Gailus-Durner V, Hans W, Hölzlwimmer G, Javaheri A, Kalaydjiev S, Kallnik M, Kling E, Kunder S, Mossbrugger I, Naton B, Racz I, Rathkolb B, Rozman J, Schrewe A, Busch DH, Graw J, Ivandic B, Klingenspor M, Klopstock T, Ollert M, Quintanilla-Martinez L, Schulz H, Wolf E, Würst W, Zimmer A, Fisher SE, Morgenstern R, Arendt T, de Angelis MH, Fischer J, Schwarz J, Pääbo S. 2009. A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell*, 137:961-71
- Enard W, Przeworski M, Fisher SE, Lai CS, Wiebe V, Kitano T, Monaco AP, Pääbo S. 2002. Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*, 418:869-72
- Fisher SE. 2006. Tangled webs: tracing the connections between genes and cognition. *Cognition*, 101:270-97
- Fisher SE. 2013. Building bridges between genes, brains and language. In JJ Bolhuis & M Everaert (Eds) *Birdsong, speech and language: Exploring the evolution of mind and brain* (pp. 425-454). Cambridge, MA: MIT Press
- Fisher SE, Marcus GF. 2006. The eloquent ape: genes, brains and the evolution of language. *Nat Rev Genet*, 7:9-20
- Fisher SE, Scharff C. 2009. FOXP2 as a molecular window into speech and language. *Trends Genet*, 25:166-77
- Fisher SE, Vargha-Khadem F, Watkins KE, Monaco AP, Pembrey ME. 1998. Localisation of a gene implicated in a severe speech and language disorder. *Nat Genet*, 18:168-70
- French CA, Jin X, Campbell TG, Gerfen E, Groszer M, Fisher SE, Costa RM. 2012. An aetiological Foxp2 mutation causes aberrant striatal activity and alters plasticity during skill learning. *Mol Psychiatry*, 17:1077-85
- Groszer M, Keays DA, Deacon RM, de Bono JP, Prasad-Mulcare S, Gaub S, Baum MG, French CA, Nicod J, Coventry JA, Enard W, Fray M, Brown SD, Nolan PM, Pääbo S, Channon KM, Costa RM, Eilers J, Ehret G, Rawlins JN, Fisher SE. 2008. Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Curr Biol*, 18:354-62
- Hoischen A, van Bon BW, Gilissen C, Arts P, van Lier B, Steehouwer M, de Vries P, de Reuver R, Wieskamp N, Mortier G, Devriendt K, Amorim MZ, Revencu N, Kidd A, Barbosa M, Turner A, Smith J, Oley C, Henderson A, Hayes IM, Thompson EM, Brunner HG, de Vries BB, Veltman JA. 2010. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat Genet*, 42:483-5
- Jarvis ED. 2004. Learned birdsong and the neurobiology of human language. *Ann N Y Acad Sci* 1016:749-77
- Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, 413:519-23
- Lai CS, Gerrelli D, Monaco AP, Fisher SE, Copp AJ. 2003. FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain*, 126:2455-62

- MacDermot KD, Bonora E, Sykes N, Coupe AM, Lai CS, Vernes SC, Vargha-Khadem F, McKenzie F, Smith RL, Monaco AP, Fisher SE. 2005. Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am J Hum Genet*, 76:1074-80
- O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S, Karakoc E, Mackenzie AP, Ng SB, Baker C, Rieder MJ, Nickerson DA, Bernier R, Fisher SE, Shendure J, Eichler EE. 2011. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet*, 43:585-9
- Petitto LA, Marentette PF. 1991. Babbling in the manual mode: evidence for the ontogeny of language. *Science*, 251:1493-6
- Senghas A, Kita S, Ozyürek A. 2004. Children creating core properties of language: evidence from an emerging sign language in Nicaragua. *Science*, 305:1779-82
- Terrace HS, Petitto LA, Sanders RJ, Bever TG. 1979. Can an ape create a sentence? *Science*, 206:891-902
- Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M, Alarcón M, Oliver PL, Davies KE, Geschwind DH, Monaco AP, Fisher SE. 2008. A functional genetic link between distinct developmental language disorders. *N Engl J Med*, 359:2337-45
- Vernes SC, Oliver PL, Spiteri E, Lockstone HE, Puliyadi R, Taylor JM, Ho J, Mombereau C, Brewer A, Lowy E, Nicod J, Groszer M, Baban D, Sahgal N, Cazier JB, Ragoussis J, Davies KE, Geschwind DH, Fisher SE. 2011. Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. *PLoS Genet*, 7:e1002145
- Vernes SC, Spiteri E, Nicod J, Groszer M, Taylor JM, Davies KE, Geschwind DH, Fisher SE. 2007. High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. *Am J Hum Genet*, 81:1232-50
- Watkins KE, Dronkers NF, Vargha-Khadem F. 2002a. Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain*, 125:452-64
- Watkins KE, Vargha-Khadem F, Ashburner J, Passingham RE, Connelly A, Friston KJ, Frackowiak RS, Mishkin M, Gadian DG. 2002b. MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain*, 125:465-78
- Watson J. 1990. The human genome project: past, present and future. *Science*, 248:44-9

