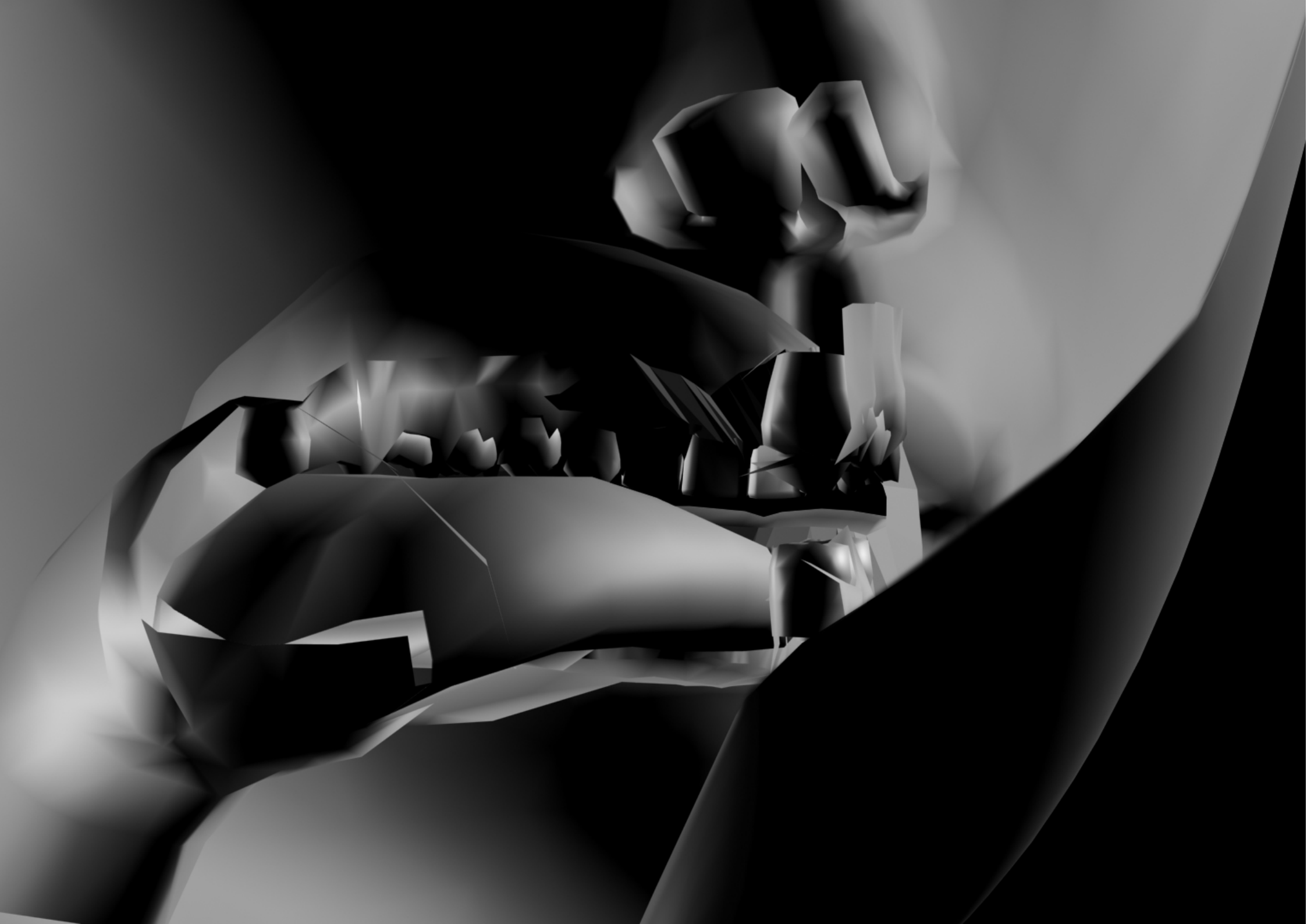

20.02.13 Artificial Voices -
discussion between Roger K Moore & Ranjan Sen
as part of 'Suppose I call a man a horse, or a horse a man?'
Anna Barham's residency at Site Gallery, Sheffield, 09.02 - 23.03.13

Anna Barham:
Thank you everyone for coming. I'm really thrilled tonight to welcome Roger Moore, professor of spoken language processing at Sheffield University, and Ranjan Sen who's a researcher in linguistics, also at Sheffield University. I'm going to give a little bit of an introduction into why I saw a link and was interested in having a conversation between the two of them.

The piece through there… I think it's just been turned off but… *Liquid Consonant** is a piece I made from a reading of Plato's text *Cratylus* which deals with language and whether it has any intrinsic meaning or whether it just operates by convention. There's a passage where he conjectures that if language really has intrinsic meaning then perhaps the individual letters and syllables would also have meaning and they would encode – he doesn't mean it in an onomatopoeic sense – but that they would encode reality and somehow mimic it. One of the examples he gives of that is the rolled r of the Greek letter 'rho'. He says that because your tongue is most agitated and least at rest in its pronunciation, it is a good tool for copying motion. He then identifies a whole series of words that include it that are related to motion – words like current and flow... I wanted to do something with that passage and it brought up the question of what ancient Greek sounded like – I was reading these words in a foreign language, I didn't know how to interpret them. In the end I worked from a video of someone pronouncing the words in modern Greek but then substituted those sounds with sounds that were more related to what the words meant. But it had brought up the idea of constructing a sound out of nothing in a way – out of a written format.

I'm going to ask Roger and Ranjan to explain or give a little introduction on each of their areas of research, but the link that I saw was that Roger, who is involved in speech synthesis, so creating voices for computers and robots – that seems to be one way of creating a voice without using this vocal apparatus [pointing to mouth and throat] and then part of Ranjan's research is into reconstructing the sounds of dead and lost languages, from written sources of course, so that seemed like another sense of creating an artificial voice.

They haven't met before this evening and they're working in separate fields and separate departments within the university, but I felt that there was some interesting overlap – or not – but certainly an interesting conversation to get them together to talk about their research. They're both going to give an introduction and then I'm going to ask some questions to turn it into a three-

* http://annabarham.net/video/liquidconsonant.html

way conversation and then towards the end we'll invite questions.

Roger Moore:
Thank you Anna, it's great to be invited to participate in an event like this, I'm so used to much more formal situations so it was quite an exciting prospect to just be thrown into a discussion and I'm very interested to see how it turns out.

I've been working with speech as a signal for many years. Originally I was running a research lab, the government actually had a centre for speech research and I was running that for a few years. Then I moved into a small company where we were trying to sell speech products – these were devices that could respond to your voice or generate speech. I realised that was not to my taste so I've ended up in academia. I've also got various affiliations; I'm currently in the computer science department but I'm not a computer scientist; I'm associated with the phonetics and linguistics department at UCL but I'm not a phonetician or a linguist; I'm associated with a robotics lab in Bristol, but I'm not a roboticist. I'm actually an engineer by training so the bottom line for me is can I create a functionality, can I make something which does something useful and interesting? That all sounds very practical and market orientated but I'm also quite well known in my field for worrying a lot about the human process – what it is that human beings are doing in recognising what other people say, in engaging in conversations and speaking?

For many years I was the person who was brought out to argue that we should bridge the gap between speech technology – that's all the practical things you might want to do with speech, and all the phoneticians and linguists and all those kinds of expertise that I know some of you here represent. On the basis that – if any of you have had any experience with trying to use a speech recognition device, maybe on your phone or you've called up some service and found yourself faced with an automated voice service – you probably found that the experience was not to your liking. And we can talk at length about that if you like – there are some serious issues. Whereas people are fantastic at engaging in speech based interaction and they can not only communicate quite exotic ideas but they can achieve fantastic things cooperatively by using their voice and they can do all this in the most awful acoustic environments – noisy stations etc. The technology that I'm involved with can't do that, it just falls over.

So people often ask me how we can bridge this gap and when I arrived in academia I thought here's a chance to really look seriously at this, is this a gap which has to be bridged? To cut a long story short, my current position is 'no'. In fact the linguists and the phoneticians, the people working on, and interested in speech from a non-technological perspective, are as equally baffled by some of these questions as we technologists are – this is my claim,

something we can debate! However there are some very interesting things going on in other fields, not obviously related to speech; in neurosciences; in cognition; in robotics; and it's those ideas that I've been trying to draw together in the last few years. That's why I'm interested in things like robots – not to make a robot which is going to come in and serve us all drinks after it's asked us what we want – but as an experimental paradigm to investigate how it is that sophisticated organisms like this one, and this one, and this one, [pointing to members of the audience] manage to do amazing things just by vibrating air between us. That's probably long enough to give you a feel of where I'm coming from.

Ranjan Sen:
Thank you. I'm Ranjan Sen, I'm a lecturer in linguistics based in the school of English in the University of Sheffield although I've also been in linguistics departments, modern language departments, classics departments, and that tells you something about the role of the linguist in modern day universities – we've fallen between the gaps in a lot of places, because we're interested in language and of course we're all interested in language, we all have something to say about language. My interest started from doing classics, so Anna's work starting from Plato is very interesting to me. I started off as a classicist looking at Latin and Greek linguistics and from there into Indo-European reconstruction looking at the parent language of all these branches which resulted in me looking at Sanskrit from the Indic branch, the Germanic branch, the Celtic branch and things like that, trying to reconstruct backwards what the sounds and the other aspects of language were of the parent language that we call Proto-Indo-European.

In order to work backwards you have to understand how things go forwards: How does sound-change work? How do sounds change over time? How does language more generally change over time? The two are bound up together quite closely. And I found that as we were doing etymological reconstructions quite often we'd have to decide on what was a plausible sound change, a plausible way in which we can account for a word in Latin or Greek or later on as a change from an earlier source. And this notion of phonetic morphological plausibility got me much more interested in how sounds are represented in the mind and how sounds are processed in the mind of an individual, as well as how languages change over time, and how speech communities change how they sound over time. So that's how I changed from being a comparative philologist doing Indo-European reconstruction into more of a theoretical phonologist looking at how sounds are manipulated and stored and structured in an individual in the mind. Human beings are very structured individuals – we categorise, we put things in categories – and we do these things with sounds as well. Sound is a very physical, continuous entity but what do we do with them? We think of them as this sound or that sound and put them in boxes.

So that's the study of phonology, but as I was going along I realised that in order to understand phonology better I needed to understand phonetics better as well. The physical aspect of how these sounds are articulated and the acoustics… what they sound like. That resulted in me doing further investigation into the phonetics of sounds and using that sort of evidence in the laboratory in order to reconstruct how sounds might have changed millennia ago – our articulators haven't changed, we haven't evolved that much in 2000 years, and our minds haven't changed that much either. Then furthermore since coming to Sheffield I realised that in order to understand both diachronic phonology – how sounds change over time – and synchronic phonology – how sounds are ordered and structured in the mind at a given time – we need to understand not only phonetics but also the psychology of language, the area of psycholinguistics – what actually happens in real time. What are the processes in real time when we're speaking, when we're listening, when we're reading, when we're writing? ...all these processes that use language. And in order to understand the psychology of language – I'm glad that Roger seemed to have come to these areas from a different approach – we need to understand how language works in the brain and neuroscience has a lot to say about things like that. Not only in humans but in other species as well. Although that's not directly related to my research I'm rapidly coming to realise that in order to understand sounds and the way sounds change, we need not only to understand the phonetics but the psychology, the neuroscience, the acquisition elements – how children come to acquire this incredibly complex system – and when things go wrong, language disorders, things like that. So that's where I've reached at the moment but my main interest has always been sound-change.

AB:
So that brings me to my first question which began by thinking about your work Roger but I think is translatable into yours as well Ranjan, which is about… I mean my understanding is that you would construct a computer programme in order to synthesise speech and I wondered if you could say a bit about what kind of methods you would use... because I know that when we spoke before you said that older models were more mechanically based – looking at how the voice physically works physically and anatomically – and then newer ones are more statistically structured.

RM:
You've just explained it! [laughter]

AB:
I'll move onto my question then… so in that case in this interplay between how you think a mind works and how successful you think a computer programme

is, do you model the computer on the brain or do you get insight about the brain from the computer programmes? If you could say something about that…

RM:
That was the big jump! So let me say a few words about the first part, and before I do that it's probably worth being absolutely clear to differentiate between devices which recognise speech, which respond to the human voice, which recognise what people are saying, as opposed to devices which speak, which are programmed to speak. But of course if we're talking about a machine to do this then it's likely to be a computer and if it's a computer it has to be programmed. But you don't just sit there and write a programme – the bottom line is do we have a model or an algorithm that we're going to implement in a computer programme? So the fact that it's a computer and the fact that it's programmed is not so important as what is the model that you're creating and that you're simulating within the computer.

As you rightly said, 50 years ago now, some of the earliest models of talking machines – actually if you go back 200 years the oldest model of a talking machine was made of wood and brass and leather bits and it was a mechanical system which you learned to operate physically – but back in the 50s and 60s when computers were beginning to come in, the most obvious thing to do was to make models of the vocal apparatus. That means models of vocal chord vibration which is the main sound generator, models of the vocal tract, which is filtering that sound and shaping it – giving it timbre, the quality of sound that we perceive in different phonetic sounds and all the other aspects. The physiology is complex but not outrageously so, and it's relatively straightforward to simulate this with huge fidelity and great detail aerodynamically – you can do lots of complicated calculations about air flow and obstructions – or you can model it at a sort of more abstract level and say well there's a sort of filtering action going on so we can just use a simple electrical analogue, and that's exactly what was done.

Some of the early talking machines were constructed that way and they… actually the early ones were virtually unintelligible. For many years people would say it was a horrible robotic voice that you could barely understand. Gradually they became more sophisticated and more understandable, and probably the most well-known example of that kind of technology is Stephen Hawking's voice which is very much along those lines but at the end of many years of development. But with his voice and with voices like that – those of you that remember the sort of Microsoft voices on your PC – had this kind of weird sound and the market was not very happy with that so then they never really made it through to commercial success.

There was a real demand for something better, something more human-like,

where the quality was good and it was recognisable as a human speaking. The solution to that was almost trivial: people realised that if you just took a huge quantity of speech and you cut and paste bits and pieces from what you happen to have and just put it back together again in the right order – there's a little bit more to it than that but pretty much it's cut and paste – you can create very high quality voices indeed. And if you're intelligent about it, about how you use the pieces, so for example if you want to say 'what service do you require?' then you get the person who's creating this big database to say that and you just take the whole sentence – that's not even really synthesis at all. But if the synthesiser has to say something that you haven't pre-recorded then you take the relevant bits and pieces, ideally large pieces, but ultimately they could be quite small pieces and put them together. Railway announcements are a version of this; anyone who's got a sat-nav system in their car using a synthesiser then that's a version of this. And interestingly that particular technique sounds very human but is measurably less intelligible than those earlier models if you put it in a noisy background, and when you listen to that [cut and paste] kind of synthesis it all sounds fantastic except there are strange breaks in it which psychologically are really damaging, whereas the older form of synthesis is smooth – it's weird but you get used to it and it doesn't have those strange breaks.

But right now we're on to a third way of doing things which as Anna said is statistical. On the speech recognition side, for many years now the techniques that were used for building programmes that will recognise what people say, have been using probability and statistics to estimate the parameters of a huge model of the way in which people speak. That model is trained on immense quantities of speech – and when I say immense… a few years ago it would have been minutes or half an hour, but current speech recognition systems of the type you might have encountered on your phone have been trained on 1000s of hours of people speaking. That's been very successful because it's defined the whole process of recognition which before was very ad hoc but which in this probabilistic framework, recognition suddenly is very straightforward to define. What you're looking for is the most likely explanation of what somebody has just said and you have all these models behind which you can then ask, how might these models have made this sound? That's been going on in speech recognition for 20 or 30 years and people have just in the last five years realised that all those principles could be applied to generating speech because in fact these models are what we call generative models. The process of recognition is essentially like saying I have a synthesiser here and I have some unknown speech, how would I configure the synthesiser to match that, and then having configured the synthesiser you know what must have been said.

Then the question is if we've got a statistical synthesiser here, what does it sound like if we listen to it? That's what's going on right now. You won't hear this out there in the real world, this is what's happening in the research labs – so

here in Sheffield and other places, and hot off the press. One of my students for example… linking in with some of the stuff I said about the cognitive basis of language... one of the things that's completely missing from the technology is the natural adjustments that people make to each other when they're engaging in spoken language. It's the kind of thing that drives the very changes that you're interested in [Ranjan] but they are ignored in the technology right now. A speech recogniser is trained on a huge quantity of speech but it is then frozen in time, it's fixed, it will do what it will do and that's that. And the same is true of speech synthesisers, whereas we all know that when people are engaged in vocal interaction they are accommodating to each other, they're learning about each other's voices, they're adopting even the words that they're hearing from another person, maybe even mimicking some of the sounds, and even more mundanely, if it's noisy they'll be speaking louder, if somebody's looking like they don't understand they'll speak more clearly, and there's all this adaptation going on. Here in Sheffield, we have a speech synthesiser which is listening to itself, so as it's speaking, it's essentially thinking to itself, how am I doing? And so if you turn up some noise it changes the way it's speaking it doesn't just speak louder which is an obvious thing to do, it actually articulates more clearly and that's all within this statistical framework.

Sorry that was a long answer....

AB:
It's a brilliant answer

RM:
...but it is comprehensive.

AB:
Actually something… I wanted to ask you later Ranjan... about whether you thought that computer generated speech could be part of the language community but maybe we can…

RM:
I should throw in a punch line, because I know this is something you're interested in, which is that all these wonderful techniques for trying to produce a voice which sounds human-like is a huge mistake, we shouldn't be doing that at all. We should be making voices which sound appropriate to the artificial entity that we're trying to create. In other words, if we're creating an application where you're talking to an agent which isn't an actual person it is far better if it doesn't sound like an actual person. The minute you hear a human sounding voice you are fooled into thinking that this is a real human at the other end, and

probably many of you have had that experience when you've called up some service and you don't realise for a little while, wait a minute this is a machine. And then you don't know what to do. So in terms of just human factors that's a really bad idea. But this is just me now – a lone voice in a reasonably big field – most of my colleagues don't agree with this position and there is huge pressure to come up with the most human voice you can and yet we don't have anything behind that to back it up. We don't have human level intelligence to tell it what to say, we don't have human level cognitive ability to interpret what the person says to it, all that's missing, so actually if you're going to create a robot which does something useful and uses the voice which is a good thing to do because your hands and eyes might be busy, then maybe that robot should speak with a robotic foreign accent. Something which tells you immediately that it has not got a full level of artificial intelligence. So I've been working a lot with voices to make them sound robotic – all my colleagues make them sound really good and then I make them bad.

AB:
You trash them!
I definitely want to come back to that, thank you, but I wanted to ask you [Ranjan] to explain a little bit about your methodologies, because I realised that when we spoke I had this picture of you examining old texts and doing it in a very kind of… in the library and then I wondered whether there was a processing element, a computer element to it?

RS:
There is to the broader question, but specifically to how I do it there isn't – I suppose you could say I look at dusty old books but in a white coat. I'm developing techniques to reconstruct fine grain phonetics of ancient languages and dead languages, non-current forms of languages. One might ask why should I bother doing this aside from just general interest? Well, in order to understand human beings, and how human beings use sound, I need to understand how sounds have been used by any language. Language is a product of the human mind, and one of the main aims for any linguist is trying to work out what language is exactly. Is language special?

Well yes it is pretty special. There's a famous quote from Bloomfield in the 1930s saying it's the most remarkable achievement any of us are ever going to make and yet this hugely complex system is acquired by every one of us. Linguists try to understand what exactly that is and phonologists try to understand how the sound element of that works. So in order to understand how sounds are structured in the mind we've got to understand any language that's been the product of the human mind and that of course is why it's incredibly important to record and document endangered languages at the moment. We can't just

look at English and Spanish and French and things like that just because we have easy access to them. All these other languages all over the world that are dying out are also products of the human mind.

That's the synchronic angle, but the same goes for the diachronic angle. What have languages in the past done? How have they behaved? What can we reconstruct about their sound patterns, how did they use sounds? This is what's led me to say we can unlock dead languages if we have a much better technique to read the data from dead languages, read the evidence and see what they're telling us. Understanding the sounds of dead languages has been going on for decades, people have been reconstructing how classical Latin sounded and things like that based on various sources of evidence, one is that the Romans themselves and the Greeks an the Sanskrit Grammarians themselves tell us a lot about how their language sounded, there's lots of detailed information like 'to make this sound that we write this way you put your tongue here and it sounds a bit like that' so they give us a bit of articulatory and acoustic information about the sound and we could relate that to the sounds of the world's languages now. That's a source of direct evidence. A source of indirect evidence is phonetic spellings: we're all used to graffiti where people say 'I woz ere' spelled w-o-z, but nice texts say w-a-s why are we not pronouncing it 'wass'? The graffiti is telling us that this word is pronounced 'woz'. And graffiti and substandard sources of writing have been used as sources of evidence to reconstruct the sounds of Latin and Greek etc. Places like Pompeii are goldmines for that, there's lots of graffiti that tells us how these sounds were made.

Of course Latin died and developed into the Romance languages so we've got an end point – we can see how the language is split up into its daughter languages, and we have quite clear evidence of how these daughter languages sound, and from what we know about how sounds can plausibly change we can plausibly reconstruct back to something of how they must have sounded. For example we know that the sound 'c' [k] in Latin has become 's' [s] in some places 'sh' [ʃ] in other places and 'ch' [tʃ] in other places, but when they've become these different sounds it's all been quite systematic. With Latin 'c' [k] has become 's' [s] before what we call the front vowels 'ee' [i:] and 'e' [e] in all the daughter languages. And from what we know about how languages change we can reconstruct back that this must have been 'c' [k] in Latin from what the grammarians tell us, from what the plausible sound changes are written consistently in Latin as well. So the romance languages give us an end point and then the Proto-Indo-European gives us a starting point.

We can compare Latin, Greek, Sanskrit and Gothic, Old English, Old Welsh, Old Irish, and Tocharian from out east – the family of Indo-European languages and using the same technique looking at how these sounds are represented graphically – as you were saying through writing – we can see how they

were represented graphically using the same technique as for Latin we can sort of work out what the same phonetic representations probably were and then work backwards from there to the Indo-European source. So from the Indo-European source we have a starting point, from the Romance languages we have and ending point, from the grammarians we have information in the middle. Of course Latin isn't one big monolithic entity, it changes throughout its history as well and inscriptions give us lots of evidence for how Latin changed. And I'm mentioning Latin but these same principles can be applied to any non-current language with similar sources of evidence.

I suppose I'm trying to go a bit further and using experimental techniques, in terms of what phoneticians have done in the laboratory with things like sounds, in order to see how sounds respond in certain environments, how speakers respond in certain environments, how speakers produce sounds and what the acoustics of those sounds are in certain environments – in systems that are plausibly similar to the systems I'm working with. Using experimental laboratory evidence to reconstruct the fine grain phonetics of Latin.

One thing I've been working on recently is trying to work out exactly how long vowels were in certain syllable structures in Latin. I'm talking about millisecond differences, but we can glean quite a lot of evidence that the Latin vowel duration was opposite to the English vowel duration and that occurs in several different languages. I'm thinking about open and closed syllables – open syllables are the ones without a consonant at the end like 'pa' and closed syllables are ones with a consonant at the end like 'pap'. In most of the world's languages, a vowel is much shorter in a closed syllable 'pap' than in the open syllable 'pa'. However in Latin there seems to be distinct evidence from how the sounds changed over time, not only in duration but also in the vowel quality, and comparing it to a handful of modern languages which have this opposite pattern as well where the vowel is actually longer in the closed syllable – Turkish, Japanese, Finnish have this pattern – everything we can reconstruct about Latin from how it changed in its development in terms of duration and quality everything seems to point to the fact that Latin does this as well. So we're using this reconstruction to build a larger theory of what sound patterns can occur in the world's languages, what influences a language to go one way rather than another way? Why do most of the languages go one way and a tiny proportion of languages go the other way?

AB:
And I wondered then… to join where you both ended up a little bit… and sorry I'm basing a lot of this on what we spoke about before… but you [Ranjan] mentioned that in fact we only use a very small group of sounds in language – that we can do all sorts of things with our mouths but we don't use them in language… and so I think that relates to this idea of what a robot might sound

like – involving other sounds that don't belong to this group of language, or human-language, sounds. Because when you [Roger] first spoke to me about the ideal sound for a robot, I just kept thinking if it spoke words how could it avoid sounding human? I mean Stephen Hawking's voice to me sounds like it's coming through a filter but I still feel like there's a human behind it. And I wonder if that links with this idea that there's a sort of subset of sounds that belong to language?

RS:
You're referring to the fact that we can make lots of different noises with our articulators I mean [blows a raspberry] is not a speech sound in any of the world's languages although children make it and it's a very common sound used from a very early age. So why isn't [blows a raspberry] in any of the world's languages as a speech sound? Well it just isn't – I can do all sorts of things with my articulators, I can [twists tongue to verticle] ake y ongue er icle an ou ike a, [unfurls tongue] but that's done by any of the world's languages. This is the logical possible number of things we do [gestures a large volume with arms outstretched] and this is what we actually do [gestures a small volume with hands nearly together]. Now this has been used for one of the arguments that language is special, meaning that there is something that we're born with, something that's innate, that guides what we can do and what we can't do. The opposite approach – the non-nativist approach - would argue that this is just the result of common sound evolution rather than human evolution. If we start off with something, whatever that something is, it can only develop it in certain ways.

RM:
That's very interesting because we've been doing some work on very similar topic because from my perspective one of the main constraining parameters if you like for the sound system is to do with the energy that it takes to get the articulators in position. This is much discussed in the literature but very little experimental work has been done to test that out. The principle is that you move… they're not big articulators compared to waving your arms and legs around, but nevertheless there's significant muscle involved in the tongue and in getting it high in front and tense it actually takes a reasonable amount of energy, and so sounds that would require that movement may be less favoured than ones where the tongue is more relaxed which take less energy. Now [begins mumbling] takes no energy at all but you'll notice that you become unintelligible so there is a balance between intelligibility – communication – and energetics. There's a guy that published a paper on this in the 70s I think – Lindblom?

RS:
Lindblom, yes.

RM:
A theory called H&H: Hyper and Hypospeech. Hyperspeech is hyper-articulated, exceptionally clear speech which takes lots of energy – which is why as a lecturer I'm exhausted at the end of the lecture – and hypospeech which is what I was doing a minute ago [mumbled]. And in order to test his theory the best he could do was to put someone's head in a box and measure the oxygen uptake – it was pretty crude. So we built the world's only animatronic tongue and vocal tract. If you search youtube* for Anton (ANimatronic TONgue) you can find a robot that we have just up the road here. If you look at Anna's work [*Liquid Consonant*] next door you see a 3d model of the inner workings of a vocal tract, but ours is a physical model. The muscles are innervated by fishing lines in fact which run to servo motors and the whole thing's under computer control. That's the first time I think to actually explicitly measure the energy involved in getting the articulators into particular positions, precisely because this is thought to be a strong conditioner on the way in which sound is organised.

Image of Anton - Animatronic tongue and vocal tract model created by Robin Hofe, http://staffwww.dcs.shef.ac.uk/people/R.Hofe/anton/tongue.html

* http://www.youtube.com/watch?v=ZFT9B6DT6wA

RS:
Speakers are fundamentally lazy and listeners are fundamentally demanding.

RM:
Yes… what interests me about that is that it's not something special in speech, it is true of all living systems. All living systems are optimising energy out versus energy in and what they need to survive, so I think it's a much broader issue that's being tapped into.

RS:
I agree that there is this dichotomy between speaker and listener but there are curious findings in linguistics where there's been articulation without acoustic effect or acoustic effect without extra articulation and things like that – acoustics and articulation don't seem to match in the way the Lindblom model would have suggested, once again suggesting that the brain is involved.

RM:
It's not just about muscle activity… there are certain configurations in the vocal tract where a very small movement has a huge acoustic effect so that would

be of great advantage and there are other regions where a large amount of movement doesn't really change the acoustic effect that much at all so why would you do that…. that kind of landscape, the energetic landscape, is the very thing which over time populations are researching.

AB:
You [Ranjan] were talking about… is it here, and somewhere near here, where the 'r's and the 'l's are kind of reversed?

RS:
Paul's the expert here… do you want me to relate that to…

AB:
Yes, I suppose my question is that that doesn't seem to be about using energy, that it's about differentiation…

RS:
Yes, in order to speak we have to make our speech sounds different from each other if we're going to maintain a contrast, and one thing that happens over time is when two speech sounds aren't particularly different, they do end up merging which loses that difference. So different speech communities can use different techniques to implement these differences between sounds and a great example is how people use 'l's and 'r's in different dialects of English.

There is no consistent way of doing an 'l' and there's no consistent way of doing an 'r' and it seems like Leeds speakers – LLLeeds [LLL sounds almost like lull] speakers – do precisely the opposite to Newcastle speakers in how they implement the way they do the 'l' – and they do the 'r' in the opposite way. 'L's and 'r's are known as liquid consonants, and they're grouped together because of the way that they behave in the phonological structure of the world's languages. Think of consonants that appear in second position after another consonant in a word – bl and br, cl and cr and things like that. They often pattern together in the world's languages in this way and this is why they're grouped together but some languages don't have an l/r difference – a lot of east Asian languages, Korean for instance, has something that sounds a bit more like an 'l' and something that sounds a bit more like an 'r' but they're not different speech sounds, they just use whichever one according to the environment it's in, whether it's between two vowels or if it's at the start of the word or something like that.

To relate this to what we were talking about… there are different ways to

implement the speech sounds that we feel are categorical, categorically different.

AB:
So then if we can move…

RM:
To the robot voice... that's an interesting one because… the reason that a human voice sounds human is because it's coming from a human physical anatomy and with all it's particular characteristics – its particular absorptions, particular resonances – so there is a timbre to the voice which is recognisably human. But any physical set of cavities with similar resonant properties can create the same kinds of patterns and will be potentially perceivable as speech. In fact I would say that any sound is potentially perceivable as speech depending on the context – if a door creaks you can hear a name being called if you're kind of expecting it because we've got an amazing pattern recognition engine up here [points to head] trying to make sense of our environment. So the question then is how would you create voices that are appropriate to particular artefacts? A simple example is… let's say a robot is typically assumed to be metal rather than made of skin. Metal is much tauter than skin and therefore the bandwidths of the robotic mouth would be much tighter, and if we model in a computer you hear a metallic sounding voice – it's still intelligible, it's still speaking normal words, but it has a sound that is clearly not coming from a human artefact.

We had a fantastic idea – well I thought – years and years ago: we were trying to the different ways that people treat an automated system depending on whether they think they're talking to a human being or a robotic agent, and what we set up is what's called a Wizard of Oz scenario – because we didn't have a robot. We had a person providing a telephone based service, and a big switch. When it was in one position, whoever called in just spoke to this person; when it was in the other experimental position, we had doctored the voice of this agent so that they sounded robotic. The person calling in had no idea what position the switch was in – they either heard a normal voice or a doctored robotic device, and we published the results. We found huge differences in the way people reacted depending on the voice, but that's not the main point. The main point is that we thought hard about how we would do this – a lot of people were doing experiments like this at the time, this was in the late 80s, and were putting the human speech through military communication devices to create a kind of weird robotic voice which actually was quite hard to understand. But we were thinking about what would be as intelligible as a human voice but clearly could not be coming from a human vocal tract. We realised it would be very easy, with some simple processing, to doctor the signal so it appeared to have two sets of vocal chords, set slightly apart in pitch. Do you get the idea?

Humans have one set of vocal chords putting sound into the system but if you have another set then it should be perfectly intelligible but it's going to sound odd. So we made it and we listened to it and it sounded like every science fiction robot you ever heard – the people in the BBC radiophonic workshop knew this stuff way before us scientists got onto it. But we used it and it's very very effective. So the message there is actually if you want to be inspired about voices like that you only have to go into the sci-fi films and see what those engineers are doing and it's incredible – the characterisation that they get with artificial voices through manipulation of a human voice.

AB:
I did just that after we spoke – I looked up your example of Wall-E and I found my own example to set that against which was Hal from 2001. And I found that they were polar opposites because Wall-E doesn't… well he says 'Eva'…

RM:
He says it with a rusty voice

AB:
Yes, he does, but the rest of his communication is kind of squeaks. I don't know if people are familiar with this film but he's very emotional, I mean you can really tell how he feels, and he's very anthropomorphic, he has these big eyes that have expressions as well. And then Hal is the opposite because the quality of his voice and his intonation is perfect, it's spot on… computer voices in the 60s when that was made must not have sounded anything like that, and yet he does sound like a computer because there is no emotional content whatsoever, it's totally void of that. And I felt very caught… neither of them seemed appropriate to me because Hal seemed to perfectly… he had all the timbre of a human voice so that too hyper-real in a way for a computer somehow to me. There isn't that robotic quality that you're talking about.

RM:
But did you notice what happened when they switched him off? They started pulling out the circuit boards and the voice degraded.*

* http://www.youtube.com/watch?v=c8N72t7aScY

AB:
No, I just watched some clips... it's a long time since I watched the whole film.

RM:
And in the end… those of you who remember the film… he starts singing 'Daisy Daisy' and that is the first… the very first synthesiser at Bell labs in the States sang 'Daisy Daisy'. If you go onto youtube you can find the original recording so that was actually a really neat link into the real technology.*

* http://www.youtube.com/watch?v=OevgCsJmeKo

AB:
The real world…

RM:
But I always took Hal as having that rather smooth reassuring voice… again I'm sure it wasn't arbitrary, I'm sure it was very carefully chosen with the design and the voice actor… because it was an intelligent machine, capable of very high cognitive activity, so of course you would expect it to have that more human-like characteristics, whereas Wall-E is a nice little rusty character... running about clearing up the garbage and talking just like you'd expect something like that to talk. Actually in Wall-E if you know that film, it's not so much Wall-E that's interesting as Eva. Eva is this sleek female, very advanced robot – even her different parts don't connect they just hover near to each other, everything is very smooth – when she speaks, it's not with a typical female voice, it's the voice of a power station and that alerts you to something which is then revealed because very early on she shoots up the entire landscape and you realise that this little egg-shape – and you already knew it from the voice – behind that is immense power, and that was just done with the voice. It's very clever.

RS:
This ties in with ideas of language and identity and what certain sounds, accents and uses of language tell us about the individual, for example if you listen to an RP speaker and listen to a Glaswegian, listen to an Irishman and then you're asked which one of these people is the most reliable or which is the most friendly, something like that… this probably ties in to something you want to ask later… but rather than it being anything intrinsically about the sounds and how the sounds are used it's tapping into our opinions on the sounds which are very much conditioned by how we've grown up and what different sounds do and things like that.

But there seem to be some universal elements, like use of intonation – Hal has a flat intonation – I don't remember the film that well – but flatter than you'd expect from a human being.

RM:
He's in control, it's important for the story.

RS:
Wall-E has very exaggerated intonation. Not only do you, when you're listening to a voice, tap into all these connections and associations that we have, but there are also these more universal elements. Well for one you're a human being and you must use some sort of intonation.

AB:
With what you [Roger] were saying about computer or robot sounds or a door squeaking and you could hear a name – I was thinking about your work Ranjan, and maybe it's speculation… maybe it's not… whether you thought that there might have been other sounds in language that have just died out. Or whether we really have just been recycling a pretty tight pool since people started speaking. And I was thinking about click languages and things like that which are so geographically isolated...

RS:
It's a very good question and from everything that we can glean about ancient languages the answer is no – they use the same set of speech sounds that we use in our… the same set of variables within which you can vary your speech sounds the same sets of vowel inventories. Vowel inventories – the set of vowels you can use – are different in every language. And the consonants… things like voicing distinction, this is what I mean by properties of sounds. But all the properties of sounds and the way the properties can be manipulated and changed seem to be consistent throughout human history as far back as we can go.

There are sounds that we reconstruct where we're not quite sure of the phonetic value, and we're not quite sure of what they match to in present day speech systems but we all kind of agree that they must be like… there are things in modern languages that behave similarly or have behaved similarly. One example are the Proto-Indo-European laryngeal system: for Proto-Indo-European we reconstruct 3 sounds which have been labelled laryngeals. The great linguist and genius Ferdinand de Saussure came up with this in the late 19th century to explain patterns in the daughter languages of Indo-European, there must have been a sound in a certain position in these words that has been lost but left a trace that affected other sounds around it in some way or other. Now, linguists, phonologists, have identified that there must have been 3 such sounds which behaved in a similar way which have been referred to as first, second and third laryngeals – most imaginatively – and we can reconstruct

from their behaviour something of what they sounded like. The third laryngeal quite clearly involved the rounding of the lips [rounds lips] o – quite clearly was something like that because of the traces it left although it was lost. The second laryngeal probably involved the narrowing of the pharynx the tube at the back, back there, probably an expansion of the oral cavity, because of the effect it had on surrounding sounds… The third laryngeal was probably something that left very little trace, something like a 'h' [very slight h noise] that can be lost very easily in the world's languages but often lengthens the sounds around it. For example if you think of a word, a Proto Indian word – Brahmin – in English you just say Bra(h)min – you lengthen the a but you've lost the 'h' and that's exactly what the third laryngeal seems to have done. If you think of French, the Latin word for the falcon – falconem – the French have gone to faucon – the 'l' has been lost and changed the sound of the preceding vowel. We know something about how sounds change and their behaviour and we can reconstruct these sounds of indo European even though we're not entirely sure.

AB:
But I guess there's the possibility of oral cultures that have just passed now and you'll never have access.

RS:
Yes, that's very true and you might say that it's a complete accident that the click languages have survived in that coming from the point of view of the speaker, clicking part way through a sentence is…

AB:
labour intensive...

RS:
Well reasonably, well – we'd have to do the measurements. But yes it is one family of languages which means that they had a common ancestor, probably the clicks were in there and that language happened to survive, so who knows, maybe… that's a very good point.

RM:
But they're going to be very prominent, very perceptible, the communicative value might be high, even if the energy cost is high, I don't know… that would be my guess…

RS:
So these are languages which involve sounds like [makes various different clicks with his tongue] and things like that and they come from one area of Africa and they're all related.

RM:
So the bottom line is the anatomy, and the control, and like you [Ranjan] said right at the beginning that hasn't changed in a very long time.

AB:
I think it's just interesting to me that with a computer, given that you could generate any sound you liked essentially, that it's still either modelled on samples or on a mechanical reconstruction…

RM:
For something to be understood as speech it has to relate to the patterning we'd expect. Of course we could generate any sound and as I said up the conditions somebody might be inclined to perceive it as a speech sound, but without all that context to help then we need to be quite close to the patterning that people are familiar with and expect.

AB:
Maybe this is a good point to open it up to some questions...

A1:
I was quite fascinated... when you were talking about the machine sound versus the voice sound... when I phone these phone lines and I need to speak to someone I find myself screaming – if it was another person I wouldn't do that and if it sounded like a machine I wouldn't even bother, it's because it gives me that human… I feel like I'm so frustrated by it… I think one of the things about machines is they don't seem to have any empathy. I wonder if the whole branch of linguistic… bionicals would do better focussing on empathy rather than words.

RM:
All of those are very valid points because your experience is not unique, this is a very common thing. There's been a huge amount of work in the last 5 maybe 6 years on the emotion in the voice and coming up with technologies which can detect whether people are getting angry on the other end of the line, purely to

catch customers who are losing it and to switch them very quickly onto a real person to sort things out. So there's a lot of very practical applications.

A1:
That would be something else though wouldn't it, that would just be picking up….

RM:
You're talking about something a bit deeper, yes.

RS:
It's the absence of real-world knowledge isn't it.

A2:
Something that always seems to be missing in these conversations is body language – so much of our communication is through visual body language and I wonder how that whole area can be incorporated.

RS:
You're absolutely right, we've been focussing on sounds and sounds are one tiny aspect of speech. You have the word construction; the sentence construction; the meanings; the pragmatics of the situation; the real world knowledge and the extra-linguistic cues and signs; what you're looking at; joint attention – there have been lots of studies with adults and children about how important joint attention; and things like turn taking – turn taking is a very human thing as well – things like that. All of these are important in language and this is why we can't just get the right acoustics we need to get the… I notice over there the sign on the door that says 'exhibition continues / push door to open' and we just know that those two statements are linked, right? We know that that means that the exhibition continues through that door. But if you read that purely semantically you can't see how they're linked, you need real-world knowledge.

A3:
You can still have a reasonably empathetic, emotional conversation on the telephone though can't you? You have timing too don't you, that's an issue. Timing, waiting to respond, and even just the 'hmm' 'uh huh', that's really helpful isn't it – 'ok' 'yeah'. But what I was going to ask earlier was this conflict that you're having with your colleagues about whether a machine should have a human voice or not… but machines don't talk! Do you know what I mean?

That's counterintuitive – a machine wouldn't talk. A machine communicates some other way and so having a human voice isn't necessarily important.

RM:
Well it depends what you mean by a machine, because a machine could be a complete facsimile of you and still be a machine so… or it could just be a little box on the floor like a piece of lego. So a machine… we have to ask what kind of machine are we talking about? And a machine which is at the other end of a telephone – who knows what it looks like but it's providing a service and it's having to use a voice because it's communicating with you over the telephone.

A3:
It's interesting because I was really kept thinking about Prometheus – it's really science fiction this conversation isn't it? There's two things in Prometheus, I don't know if you've seen it but, the kind of synthetic man, David … there's two things… the first thing is that he reconstructs the languages of the visitors by putting together all the common languages and he's able to communicate with them. But the second thing is when they're all going out into the atmosphere he puts the helmet on and the guy says 'why are you putting the helmet on – you're a robot?' And he says I'm made this way to make people like you feel more comfortable, so if I don't put the helmet on it's going to freak you out. So I guess whether the thing has the voice or not is important psychologically to how we react.

RM:
That's very interesting because people do freak out when they get mixed messages, when they get confusing cues, and that's really the point about the voice because the intelligence behind an automated system right now is very low and yet it's fronted by a voice that's purporting to be from something with very high intelligence. That mismatch is the source of the problem.

A3:
I don't mind that robot voice, or a semi-human robot voice, but because I have a regional accent it makes me – speak – like – that – to compensate.

RM:
Unfortunately that will just make things worse. If you want to have a laugh about that go to youtube and search for eleven and lift* – it's a sketch about two Scotsmen stuck in a voice controlled lift, it's very funny.

* http://www.youtube.com/watch?v=sAz_UvnUeuU

A4:
I'm interested in your probabilistic models, because what it is you do, you pick out the linguistic parts of the voice but do you use the probabilistic models to pick out the emotion or the gender or the individuality as well? Can the other aspects of voice…

RM:
Yes, you can, indeed all those aspects are model-able using probabilistic techniques.

A4:
So do you incorporate them into the robotic voices that you produce? We identify people by their voices so presumably you could identify a robot by its voice?

RM:
You could indeed yes, and your identity in your voice comes from lots of different aspects, it comes from your actual personal anatomy, but also the way you use it, and accents, dialects, your linguistic background, how fast you speak, your expressivity, whether you're a very animated person or speak in more level tones… all those things can be captured, modelled, and put back in to an artificial voice. You can find examples of things like that. One of the main suppliers of speech synthesisers is a rather nice Italian company called Loquendo and if you go to their website and look for the demo section you can type in any text you like and it will speak it in a bunch of languages, you can select languages and you can also put in emotional markers – you can make it laugh and cry.

A3:
I have a question about... you know when we were talking about energy use… your measurements of it… and the reason why we don't use crazy sounds because we've evolved to be more economic. But what about when a person becomes emotional, does that affect… does that mean you start getting un-economical with your energy. Have you measured a person getting in a rage?

RM:
It modulates because emotion… it depends again because emotion is multi-dimensional… but emotion tends to affect your whole body, it affects the physiology in particular ways, it will affect the lungs and all of that is going to have an impact on what you're trying to do here [gestures to throat and mouth].

There are some neat theories on emotion and the voice actually and one of the theories is about the ways in which people try to cover up the emotion in their voice. We're all aware of that – when you see someone who's not used to speaking in public and you can here it go [makes his voice tremble], but you can also hear that they're desperately trying to suppress that which is kind of making it more obvious… so there's all of that, all of that is going on, it's a whole body thing and you can't…

There's a professor in London University, a neuroscientist, who's doing a lot of interesting work on laughter. Laughter is a very interesting physiological response, and again it gets the whole of your body vibrating and that's... again online you can find fantastic examples of newsreaders, very serious people on radio 4 losing it because something funny has happened and you can feel that whole emotion building up, just in their voice, you're empathising – who was talking about empathy? – so you're empathising with what's going on and you know that they're desperately trying to stop this laugh that's coming and eventually there's just silence and squeaks and noises and you know that they're having real trouble controlling their physiognomy.

RS:
You wonder how things like laughter evolved. Obviously it's a very rapid intake and breathing out – perhaps getting more oxygen into the system…

RM:
It's something like that

RS:
But why did we start laughing?

RM:
Chimpanzees laugh

RS:
Do they? There are some interesting ideas about why we smile as well, how it affects the shape of the vocal tract [smiles and his voice changes] by doing this. It gives very ee like sounds, and these are the very tiny weeny ee like sounds. Rather than the big bombastic sounds, so the shape that is created by the vocal tract is… there are sounds that are associated with being nice and passive and…

RM:
There's size as well. It's possible to judge, not their height, but the size of the vocal tract from the sound of the voice, and some animals exploit this and have mobile larynxes – the red deer for example when it wants to attract a mate and appear to be a very large healthy male, the larynx, because the deer has a very long neck, the larynx drops right down when it roars so it sounds much bigger than it really is. People exploit the same thing as well…

A3:
Margaret Thatcher took lessons to lower the tone of her voice to make her more authoritative. And David Beckham!

A2:
Yes, to make him sound less girlish.

RS:
It is interesting how we can work out sounds without hearing them. By doing that [pulling a face] you can work out what effect it will have on the sound…

AB:
I think *you* can but I'm not sure that…

RM:
I think you can recognise whether people are smiling by how they sound.

AB:
Oh yes that way round, but I don't know that I would be able to think abstractly what effect a certain shape face would have on the sound…

RS:
But I'm thinking more unconsciously of how the whole thing would have evolved in the human race to start with. And then some very interesting papers * as well looking at… while we're listening, the speaking movement part of our brain is being triggered. How on earth is that? We're not moving our articulators while listening but there is obviously some relation.

* Watkins, K. E. & Paus, T. (2004). 'Modulation of motor excitability during speech perception: The role of Broca's area', Journal of Cognitive Neuroscience 16: 978-987

A3:
I thought they'd scanned babies' heads... while they're listening to their parents talking the parts of their brains for speech and language is developing and so unless you hear certain sounds by a certain age you're just going to not be able to articulate them very well at all. That's why maybe you know you were talking about the r/l thing, some people struggle when they learn another language to say the sound properly.

RS:
There's attenuation… from 0-6 months we seem to be very perceptive to a lot of the sounds of the world's languages. 0-6 month old babies can tell the difference between the Hindi consonants da da da pa – I'm exaggerating them a lot but lots of adults find it very hard and yet 0-6 month old babies can do it from an English speaking background. Which possibly suggests that we're born with some mechanism to discern between these speech sounds and then we lose it.

A2:
I taught my daughter to speak using sign language because she has a learning disability. So we learnt Makaton. It was having the word plus the gesture, so orange [gesture of the sign - raised hand as if holding an orange] and she would do the gesture before she could say the sound. But she got on and she'd start using the sound, start copying what she'd heard and then she'd just get the sound and drop the sign – without being told 'don't sign', she just dropped the sign and didn't use it anymore.

RS:
Sound and meaning in the real world

A2:
That's right and the gesture... to use a gesture alongside a sound seemed to cement the sound and the understanding of the word in the formation of learning.

RS:
That's very interesting… similar things with pointing as well. So you have to make the sound-meaning link and pointing is our way of relating it to something in the real world, but what you said is interesting because it's something symbolic rather than something in the real world.

A5:
I was wondering Ranjan about… you were talking about how in some situations you get the vast majority of the world's languages that do things in a particular way and then you get a very small subset that does it in another way… I wonder if you think we could learn anything from that sort of – if you like how Roger's doing, with probabilistic spread, that tells us that this thing is more likely than that thing, or is it really just historical accident? And if it's really just historical accident can we actually learn anything by the fact that most languages do something in one way or is that just complete chance, coincidence, do you think?

RS:
Well it's… I wasn't quite sure about your difference between probabilistic and chance, there are chance happenings because of the likelihood of something happening…

A5:
Well if you have 95% of the world's languages that do thing x and 5% do thing y and you're trying to model on that and say let's do x, but actually you ignore the fact that some of them do y.

RS:
Everything you're assuming has started off in this pool of phonetic variability, every speaker speaks differently, every speaker speaks differently at different times, in different circumstances… and so there's a lot of phonetic variability in what we do, and sounds vary in certain ways more than they vary in other ways and phoneticians understand how they vary. I suppose over time some things are more likely to get re-analysed in one way rather than another way, given how this pool of phonetic variability works. The sound [k] is more likely to get analysed as [s] as it has been in the romance languages because we know that the [k] before the vowel 'i' is more front and things like that but it won't go the other way because in no environment is a [s] going to sound like a [k]. But I'm veering away from your question…

I guess if we're going to put this in a model then you need to know what your phonetic variability is and then consider how that phonetic space can be divided up and decide on the likelihood of the divisions given the system as a whole – what other consonants are there in the language and what contrasts you have to form, what vowels are there in the language and how different you need to keep them apart… things like that. I don't know if that answers you question really...

A5:
I was sort of also thinking about something you mentioned about endangered languages and that there may be things in endangered languages that although they're exotic in the sense that they're not common in other languages, but if we don't find out about how endangered languages work then we don't understand some of the possibility of what we can do as humans. Of course there are odd exotic things in well established languages like 'th' in English is weird, it's a very very strange sound in language generally but it just so happens that this hugely politically influential language, because of American and British military and economic might, has become a dominant language in the world.

RS:
Absolutely, this might be an illusion that these are all the possible things that we can do [gestures large volume with arms outstretched] and that's what we actually do [gestures small volume with hands close together] – I was talking about sounds but that same thing goes for sentence structure and word structure and things like that. I mean out of all the possible ways you could structure sentences, human languages use a very tiny proportion of those – why is that? But yes looking at endangered languages and reconstructing what former languages have done might suggest that that's actually an illusion and more things are possible.

A5:
Maybe the things that we don't do are accidental…

RS:
Right but the question always comes in: are they accidental because of some communicative function? Is there something in the communication – at the end of the day we're trying to get a message across not logically form a sentence in some way – is there something about the way we communicate that suggests we shouldn't form our sentences like this, we should do it more like that? The problem with linguistics is that quite often these problems have been looked at in isolation – like sentence structure and sounds and things like that – whereas we actually have to look at languages as a whole and consider the whole system.

A1:
There's a word in Turkish for flour or corn - it's 'unh'. That's it, it's just that, and every time I hear it or see it written, it just feels like such an old word as if it goes right back to living in caves. 'unh' - to describe bread basically. And I just wondered about this pool of word resource, the further back that you get they get less and less… you know 'cave', 'fire'… they didn't say 'prefix'….

RS:
Within the realms of what we can actually reconstruct, so I'm talking to about 4000BC, that sort of time period, in terms of Indo-European languages. In that time period obviously we can't reconstruct an enormous vocabulary because we can only reconstruct things that have come down to us but some of those are very complicated words and certainly how words are constructed, how sentences are constructed and things like that are just as complicated as modern languages as far as we can tell.

A1:
How does that translate to the future? We've got a problem where we're digging holes and putting stuff in it that's dangerous for 250,000 years and somehow we've got to inform the people that come after us that it's dangerous. Surely it's got to be an oral tradition 'that's dangerous', how does 'nuclear waste' as a word evolve over the next 250,000 years. We can't draw a stick man because there's no guarantee that they'll have the same number of arms and legs in 250,000 years [laughter] Is there a possibility of an oral tradition or are we going to lose something really important that we need?

RS:
Well the oral tradition is fundamentally what language is… if we pass things down through sounds from generation to generation… writing is a construct, writing is a product of our histories, our cultures and things like that, you only have to look at how we spell English words to know that it's only about 500 years old and it's a product of our cultural history, it's a construct. So the oral tradition that fundamentally is what language is…

A6:
Anna you might know this, or someone else might know but I vaguely remember a reference from Rabelais' Gargantua and Pantagruel about where new words come from – that they're frozen in the north and when they're needed they melt in people's mouths. Does that ring a bell?

AB:
No, I don't know that, it's very nice.

A6:
It's just a beautiful passage about the generation of words melting from ice. I'll look it up… it's definitely Rabelais.

AB:
Maybe that's a nice place to wind it up. Has anyone else got a burning question that I'm cutting off? Then thank you very much Roger and Ranjan, it's been brilliant, and thank you to everyone for all for your contributions. Thank you.

With thanks to all the speakers

Links:

Roger K Moore
http://www.dcs.shef.ac.uk/~roger/

Ranjan Sen
http://www.shef.ac.uk/english/people/sen

Suppose I call a man a horse, or a horse a man?
http://supposeicall.blogspot.co.uk

Anna Barham
http://www.annabarham.net

Site Gallery, Sheffield
http://www.sitegallery.org/

LOTTERY FUNDED
Supported using public funding by
ARTS COUNCIL ENGLAND