

Confessions of a Phonetician

Fred Cummins
UCD School of Computer Science
University College Dublin
fred.cummins@ucd.ie

January 20, 2025

1 Dedication

This piece of writing is dedicated to, and written for, my many friends who have contributed whole universes of thought to my lost fumbling attempts to inquire into things. It is a thank you note, a love letter, and a lighthearted survey of some things that we have collectively brought forth. There is a path from training to wonder, from science to bemusement, and from English to mantra that happened to be the path I wandered along, but the path is constructed exclusively of the inquiry, the distributed inquiry of uncountable souls looking into things most generally, and though I do not have a map, I think the spirit of the inquiry, which belongs to all of us, is very reassuring.

2 Hunting voice onset time

To be a phonetician is to be one of the motly group of specialists who study language, I suppose. I first encountered phonetics as an undergraduate student in a laboratory in Trinity College Dublin, where we used computers with expensive software to measure certain things that were believed to be of interest in the speech waveform.

What is a waveform? A voice sounds, and causes the membrane on a microphone to flutter. Physical measurement readily produces a picture such as in this figure:

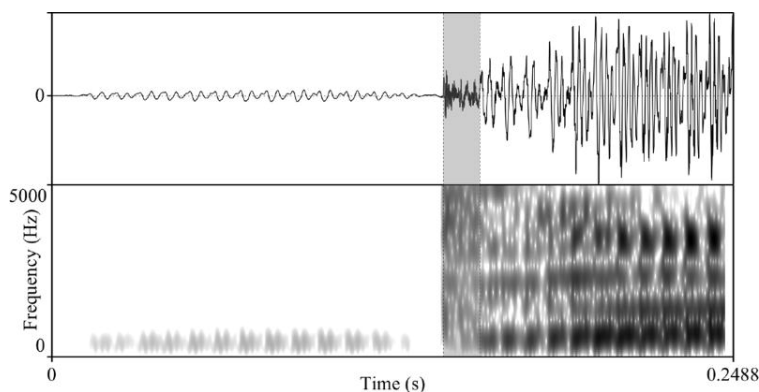


Figure 1: Screen display showing what we looked at as we sought the elusive “voice onset time.”

Ignoring all details, we were supposed to be measuring the width of the area of the gray box, and this was called “voice onset time” and it was considered important in the study of languages. I thought it rather boring and I determined not to study phonetics. But that was not what happened. The idea of the voice onset time turned out to be more important than anything in the universe, because phonetics is the fundamental science of creation for an embodied being. But to see that, I had a path to stumble along, ignorant of what I was about, looking for whatever my febrile mind thought a junior scientist should look for (and very certainly aware that I did not know what I was looking for).

3 Big for a duck

I entered a graduate programme in cognitive science in Indiana University, where a colourful array of classes was available on diverse and fascinating topics. There was excitement in the air about some new mathematics called non-linear dynamics which was palpably spreading across laboratories from its birthplace in the United States. There were brilliant professors and students, juggling, carousing and I had a most wonderful time. But my new professor, Bob Port, was a phonetician, skilled in the dark art of “voice onset time” and other phonetic things. The other things he played with were fun, including waving a slinky toy around to see it wobble in a mathematically satisfying manner, and I was captivated. He was after the music of the voice, its rhythms and melodies, its roughness and its malleability. This was speech, and there were waveforms, but we were after the pulse of the voice, the beat. To find the pulse, I had friends repeat short sequences of syllables such as “Big for a duck” into microphones while they tried to speak in time with the tic-tock of a metronome whose timing I controlled, and we had the satisfaction of being able to show other people that we could demonstrate what everyone already knows, that there is rhythm in speaking. This is frequently what happens in science: a great deal of effort is expended to show something every fool already knows. Only then it is baptized “objective” (and killed stone dead). But there is wonderful sport in the quest to show something that one has brought forth oneself and that makes a common truth obvious, manifest. It was always the scientific figures, the pictures, that fired the imagination up the most. I am certainly not alone in that.

4 Coupled heads in a dream

A year or so after I earned my doctorate, I was working in my first research position and now there was no denying it. I was regarded by others as a phonetician; I did phonetician kinds of things. I just didn’t measure voice onset time. That was for squares. I had no clue what I was looking for, though. My doctoral research had more or less come to an end and I was measuring other phonetic kinds of things without any great enthusiasm or sense of purpose. One lunchtime, just after putting my feet up for a nap, I drew a little sketch which is shown here:

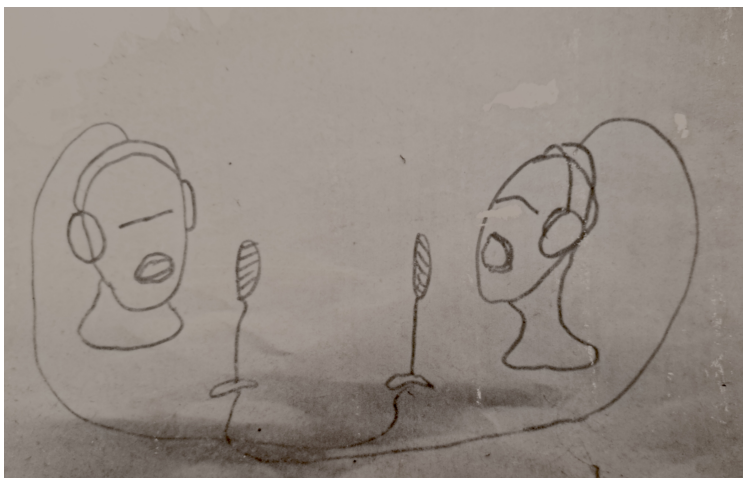


Figure 2: A sketch of two coupled speakers, done just after taking a lunchtime nap.

I had long been working with people speaking into microphones, with metronomes, listening to things I played them, ticking boxes, measuring things on computers, and it all seemed a bit dead. I don't know what I dreamed during that nap, but two years later, and in a new job as a professor myself, I came across the sketch. It was a good time to find it. I had even less of an idea of what I was supposed to be looking for: I had spent a year in an artificial intelligence laboratory and that was fun, but technical. It lacked the sonorous passionate magic of the voice. My new job was in a computer science department, but it was a nice place and I could research (that is always the word they use) pretty much anything I wanted. The trouble is, I now knew for sure that I did not know what I was supposed to be looking for, or what research was, and I was a bit stuck.

And so these two coupled heads provided me with some sport. I started simply looking at people speaking in time with other people, instead of asking them to speak along with beeps, lights, or any other controls. All I had to do was give each a microphone, and a short text to be read, and I would count down "three...two...one..." and they would start speaking the words of the text. I must have thousands of such recordings, because it turned out to be jolly good fun. People are brilliant, and they did not mind providing their speech for me. And there was lots to find out, that is to say, there were things that were not obvious but that were interesting. But for me, from the start, the wonder of the thing was that they could do what I asked at all. It was wonderfully easy, and their voices were astonishingly exactly aligned, and they did not need any practice. It all felt more real than the kind of speaking I had engaged with before.

Here is an example of the kind of text we would have them read together:

When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colours. These take the shape of a long round arch, with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow.

I'm still very fond of this little piece of doggerel. It ain't poetry, but it is a standard text used, for example, as a reading sample for deep sea divers as they return to the surface, to check for any

signs of decompression sickness.

5 Synchronous speech

The laboratory ruse of having people speak in synchrony gave me a great deal to think about, and I found a few things while playing with it. Some things are easy to measure on a computer, and so I took to measuring time on the waveform display again, not to look for the dreaded “voice onset time” but to see how far apart the speakers appeared to be. At that time, I was happy to report quantitative estimates of a distance between the speakers, and the numbers were small numbers, small numbers of milliseconds, which assured me of what I (again) already knew: they were very tightly coordinated.

I played with people who were, for me, speakers, and from their point of view I appeared to be running experiments, so they were subjects. I think it important to try to recognize how we see each other as we do scientific work. I had them face each other, face away from each other, speak with recordings of others, speak with recordings of others who were themselves recorded while someone spoke together with a person, I had them read cunningly rhythmic and arrhythmic texts (“Big dinosaurs and bigger Daleks in battle” is one I was fond of), I recorded them in groups of 4, in groups of 8, I played distracting things at them. I had no clue what I was looking for, but I kept getting numbers the profession was happy to publish.

I was only finding out what I already knew: they were very tightly, exquisitely, coordinated. It was easy to do what I was asking them to do, even though what I was asking them to do was a thoroughly artificial task that I had made up because I thought two coupled heads were more interesting than a person and a clock.

The tightness of the speakers was nicely illustrated by a breakdown that occurred occasionally, and that was very idiosyncratic: Sometimes, while reading the text, one or other would falter, and they would suddenly, in mid-syllable, both stop speaking at exactly the same time. If a person you were talking to were to abruptly stop like that, you would might look around for a sniper, yet it happened many times in the laboratory. I told the profession of this, because it seemed to me to be something they would be interested in; to be specific, I spoke of it as a “speech error.” That was an error on my part, and nobody cared, and I already knew the speakers were tightly coordinated, like dancers, and nobody is surprised if two dancers become uncoordinated together. But how many speakers are there really? One, or two? We might revisit this thought.

6 Funding, and What Now?

Things were going nicely. I could play with the coupled speakers, find things, and the fact that I was only finding things that were already obvious to me meant that I could keep on doing this. But why? What was I looking for? I don’t think anybody ever knows why they expend their energies in one way rather than another. I like to think of an unspecified, always ongoing, inquiry, with an emphasis on the kind of fixated staring that the word ‘inquiry’ can imply. I was staring at something perfectly normal.

Something happened in the intervening years. Some brilliant colleagues¹ who enjoyed the idea of staring at people speaking together joined me in crafting big funded research projects. That

¹Thank you, Roger Moore, Jens Edlund, Sophie Scott, Nick Campbell, Andrea Schiavio, Chris Newell, Petra Wagner

meant engaging with funding agencies, and explaining why you want to do one kind of research rather than another, and what you expect to find, and what training you will provide for young scientists, and a month-by-month plan of what you would do, publish, and present. The trouble is, I had no idea of any of this. Nor did my friends, who all had slightly different skills and interests to me, but together we did our best. We did not get funded for a big grant, so we regrouped and tried again, and did not get that, and we regrouped and tried again, and we did not get that. Each iteration took about a year. The exact people varied, the purposes we proposed were a colourful fabric of plausible promises, I found myself pricing a monkey for use in neuroscientific research at one point (but I forget the price), and all this, too is perfectly normal.

There is great virtue to such ‘unsuccessful’ work. Every time, a bigger story had to be told. Benefits for mankind had to be invented (we were just inquiring—inquiry does not know of benefits to the mythical beast of mankind). Above all we had to make it plain that our fascination was grounded in some larger picture. And so we pulled back, and set the previous work in a bigger picture. The bigger picture came into focus as “chant.” That word kept coming up, because the nearest thing outside the laboratory to the situation of the subjects inside the laboratory was those occasions when people speak together, that is, speak or sing, or roar, common words. There was no shortage of examples. And the examples (from the ‘real’ world, remember) were always pregnant. Our attention was drawn to the heaving crowds in football stadiums, to practices of prayer and ritual, to the ever lively protests—about this time the tragically misnamed ‘Arab Spring’ was erupting across near Asia and North Africa. There were so many examples, it was hard to know what to focus on, except to insist, in each plea for funding, that what we were looking at was real, was important, was powerful. And so, in an attempt to point to this activity that was ubiquitous and apparently undescribed, I coined the term “Joint Speech.”

7 Joint Speech: early days

The coining of the term Joint Speech was important for me. Most scientists could not point to the thing they believed they were studying. If you are a linguist, please show me some language! If you present text I will point to its material being as marks on a page. If you point to the sounds I am speaking, I will still not know what you mean by language. This is by no means unique to the topic of language. Who can point to ‘perception’ or ‘memory’ so that we can agree what we are observing and interrogating? Yet joint speech is easy to point to. Pass by a primary school and look at the children. Somewhere they will be chanting nursery rhymes, or ganging up in collective bullying intoning ‘nyah nyah na nyah nyah.’

My first try to use the term joint speech to fix an object of inquiry (a fundamental move in scientific work) was to say it was “multiple people saying the same thing at the same time.” That worked fairly well, for a while. I saw great advantage to having such a definition. At conferences and talks, I did not have to spend time trying to teach people I was talking to how I saw the world, or what body of theory my work depended upon. I simply pointed. I could play examples. Who hasn’t experienced the transcendent unity of singing in a choir? (Perhaps the fact that I was not allowed to join the school choir because of my unstable adolescent voice left a grudge that I have been pursuing?)

As linguistics became more and more foreign to me, joint speech provided a touchstone I could come back to. Now, I no longer needed a laboratory. It was enough to keep my eyes and ears open. The humdrum word ‘chant’ was very useful: I could survey thousands of newspaper headlines that

used the term and see what they had in common, or what themes recurred. There was, in a way, too much to see. Joint speech was everywhere. Over time, the definition changed a little, as we will see. It is probably never going to be fixed to my satisfaction.

8 I am not alone!

Although the umbrella term “Joint Speech” was missing from the landscape I had access to, it allowed me to recognize links between some very diverse observations of others.

Jeremy Woodruff is a composer and artist from Germany. He was living in Istanbul during the 2015 attempted coup d’état, and he and others took pains to record ‘soundscapes’ of the many protests that erupted about that time.² In particular, Jeremy keenly asked if one could tell from the sound of chanting who was chanting, and the givens of Istanbul in 2016 drew out a specific contrast between Kurdish protests and supporters of Tayip Erdogan’s regime, between a bottom up grievance of the oppressed and the reinforcing and somewhat intimidating chants of the authoritarian regime. It was also a master lesson in listening. The Kurdish protesters expressed their deep seated grievances by drawing on their stock of musical and football chants. The voices were clearly both female and male. The chants were varied, and improvised. The nationalist supporters of the regime were exclusively male, and their chants drew from very little. There is something degenerate about the chant “U.S.A.” that has become familiar in the fractured American landscape—little more than a bark. Such were the chants of the nationalists. Attentive listening to chants, taking the whole context, the nature of the participants, the timing, the purposes, is rewarding, but one has to suspect that there is richness in there to attend in the first place.

Meanwhile, in London, my very good friend Sophie Scott came up with a most wonderful, if unlikely, way to look at chanting: in a brain scanner! Now there can be no worse place than a scanner to chant in. One is lying on a bier in a large and intimidating metal device that makes a lot of noise, and one is not allowed to move. No passion, no ritual, no crowd. But it is a very precise device, and it had proved interesting to look at synchronous speech in a laboratory, away from the passionate entanglements of chanting. Perhaps there was a way to use this formidable high-tech science machine to look at something. And as always with joint speech, the important question is: What do you look at? Jeremy had had to draw on all his rich knowledge of the Turkish situation, attending to everything he could possibly attend to, to spot the obvious(!) difference between the protesting groups. Sophie was particularly well placed to wield this fantastically complex, fetishistic tool of observation: I had known Sophie since we were both young phoneticians looking to find the obvious(!) rhythm in speech, and since then she had become an expert in both cognitive neuroscience and in laughter. Laughter is interesting. It fights with speech. It is very hard to laugh and speak at the same time. Sophie was a stand up comedian as well, and she had developed years of experience getting people to laugh in the brain scanner. If anyone could find a way to look at something here, it was Sophie.

Together with her PhD student Kyle Jasmin (and a suite of helpers³) a protocol was devised that allowed a specific contrast to be precisely specified between two conditions of observation. I

²<https://seismograf.org/fokus/sound-art-matters/the-audibility-of-rebellion-versus-nationalism-in-the-sound-of-political-chant>

³Jasmin, Kyle M., Carolyn McGettigan, Zarinah K. Agnew, Nadine Lavan, Oliver Josephs, Fred Cummins, and Sophie K. Scott. “Cohesion and joint speech: right hemisphere contributions to synchronized vocal production.” *Journal of Neuroscience* 36, no. 17 (2016): 4669-4680.

think it important to be clear what that contrast was, because I thought the resulting findings were very thought-provoking. It is important to know some things about what goes on in a brain scanner to recognize the contrast.

Subjects in the scanner had (metal-free) microphones and headphones, so they could speak to, and with, someone from the experimental team in an adjacent room. Although scanning requires that the person lie still and not move, they could speak a short phrase, and then we could scan afterwards. Importantly, and perhaps counterintuitively, this brain scanning technology (functional Magnetic Resonance Imaging, or fMRI) does not measure anything at all, and it does not examine the brain. It records blood flow in the head (presumably blood that is going to the brain) and bloodflow changes slowly. Thus it was possible to speak and, in the next 5 seconds or so, to record the bloodflow. Then, after doing this under contrasting conditions, we could compare bloodflow across the two conditions. Brain science is weird. After a whole lot of work, we end up with a picture (again, the pictures do all the heavy lifting) in which a coloured area overlaid on an image of the brain represents a tiny surge of blood that is present in one condition and not in the other. A tiny surge of blood. That is what we are hunting.

Details aside, the stand-out finding was this: We had a clear and stark contrast between speaking in synchrony with another live flesh-and-blood person, and speaking in synchrony with a recording of that self-same person. Care was taken that the subjects in the scanner did not know of the contrast we were drawing out. They knew nothing of recordings, and they assumed that they were always speaking with the same person. But there was a reliable and fascinating change in blood flow if the two people were in communion (the term seems warranted, and will prove helpful as we go on) but not if the liveness of the voice were merely reproduced. We could see the connection between the two in the blood flow, though the speakers themselves knew nothing of the contrast. The main finding is reproduced, with no concern for precision, in Fig. 3.

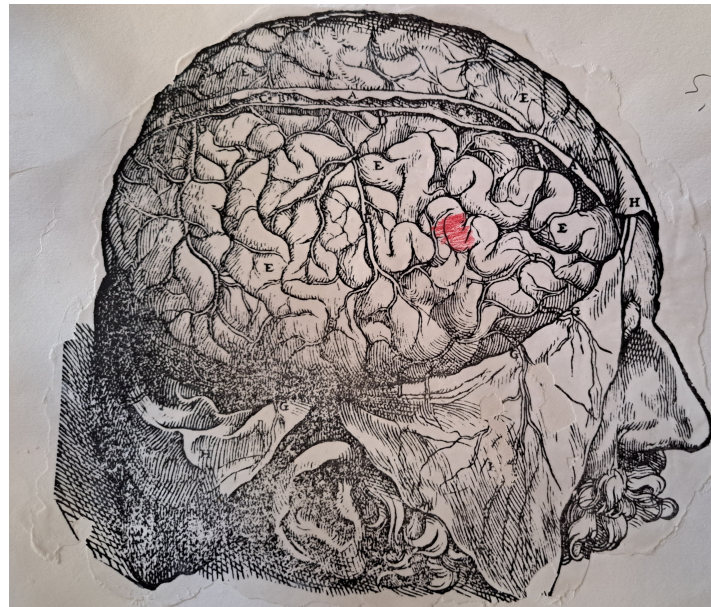


Figure 3: Rough position of an important and tiny change in blood flow found when two people speak in communion with each other. Although this is not the illustration used in the *Journal of Neuroscience*, for hermeneutical purposes, it might as well be.

What you see in such work depends on what you are asking. It is utterly conventional to set up contrasts in the scanner and look for blood flow changes. To my jaded mind, most such contrasts are uninterpretable—after all, a tiny change in bloodflow does not interpret itself. If your contrast is designed by a psychologist or a marketer, the interpretation of the tiny change will be in a psychological or an economic register. Forgive my scepticism, but I do not think the brain has any intrinsic connection to those two venerable fields. But that is merely my opinion. Our question had to do with something we found compelling: there is a liveness, and an essentially embodied character, to joint speech that is missed if we do not attend carefully.

When the 2020 Covid pandemic hit, the whole world became aware of a technological glitch that my friend Jens Edlund and I had been discussing for the previous decade: you can't chant over Zoom. You can't chant over networks. It doesn't matter if it is Skype (where we first observed this) or Zoom, or a 6G or a 7 G network. Everybody has tried to sing Happy Birthday (my favourite quotidian chant) over Zoom and found it doesn't work. It is not as if we had high standards for musical fidelity when we sing Happy Birthday! We tolerate squawks, tonal chaos, and sloppiness without problem when we are together in a room. Now here really is food for thought, and perhaps a good time to be a little self-aware of our modern belief in communication networks.

Chanting is participatory in a way that requires careful consideration. The modern condition of technological embedding makes us all, perhaps, aware that there is something precious (but not too precious) about the lively heaving body presumed to be at the center of individual experience of embodied beings. But the terms we use are confused here. I teach in classrooms where we participate richly in discussion, breathing the same air as we do so. Unfortunately this normal form of teaching acquired the wrong label during the pandemic. People spoke of this mode of teaching as face-to-face, and they distinguished it from Zoom, whereas a Zoom encounter is very well described as face-to-face communication and the classroom is most certainly not face-to-face. To put my face in front of another face in the classroom would be rude. That is not how we inhabit a room. This language confusion runs deep. We speak of our face as if it held our identity. But the face is the most abstract, socially-constructed, neurotically hallucinated thing possible. We can indeed present a face on Zoom; we can even substitute a cartoon face. It works fine. Faces are not real, in an important sense. Perhaps that is why AI is so good at recognizing them.

Something is lurking here, and I am expressing my own confusion. If you will bear with me, the confusion will get much worse shortly. I am at some pains to resist abstraction away from the body and its living. That is, after all, the basic thrust of Embodied Cognitive Science, which well describes my concerns and those of many of my friends. Cognitive science is a peculiar business. It must take on the puzzles of the knowing (what that is), knower (who can be said to know?) and the means of knowing. This way of understanding cognitive science is, perhaps, a bit abstract, but caution is required for the grand ambition of saying anything concrete here, when we are, individually, so very ignorant. Furthermore this way of outlining the general scope of a form of self-inquiry resonates with thinkers throughout many cultures and times. The joint consideration of knower, knowing, and means of knowing is found in the Theology of Proclus, in the austere tomes of Maimonides, in the contemplative figures of Tantric thought, and beyond. In embodied cognitive science, there is a shared zeal to center the body, in as many ways as possible, for there are countless ways to regard a body (as we shall see), no one of which may claim to be a finished, objectified object. To an Embodied Cognitive Science, the body has something very important to do with subjectivity and with intersubjectivity, with the magic of a shared world, despite the puzzles that this immediately throws up.

First, though, another accidental research finding! As a beginner in the dark art of phonetics, I was schooled to look for ‘invariance.’ Phoneticians are trained to get people to speak in ways such that something, anything, is invariant. A classic example is the ruse of getting people to monotonously intone sets of sentences, such as “I said heed again; I said who’d again; I said hood again; I said head again; I said had again; I said hawed again; I said hut again.” In this instance, what a phonetician is seeking is a bit odd. It is a vowel, or rather, a set of vowels, because the assumptions of phonetics include the assumption that the building blocks of language are to be found in the noises we make when speaking, and these are the vowels and consonants. Voices are unruly, though, and the ‘carrier sentence’ “I said *X* again” tries to hold the ‘context’ steady, so that the vowels can be recognized. When we speak, the speech we produce is directed to someone. We do not speak in the same manner when whispering to our elderly mother in her sick bed and when shouting at the kids to get off the lawn. The voice is plastic, malleable, and never entirely ours (again!), because it goes from us to another. The more I looked at joint speech in all its forms, the less I understood speech as relevant to letters, language, and all the usual jazz. Conceptual confusion abounds, but a phonetician who has gone rogue is still a phonetician. I found the invariance we had been looking for in our laboratories. I found it in the recitation of the rosary, kindly provided by a student friend, who had recorded her granny reciting one decade of the rosary. I took this recording seriously, and I looked at its waveform on the computer screen, and I compared one phrase after another. The Catholic rosary repeats the “Hail Mary” ten times in a row, then forms patterns out of these so-called decades. The Hail Mary itself has two halves. The first half is “Hail Mary, full of grace. Blessed art thou among women and blessed is the fruit of thy womb Jesus.” Repetition is conservative of form. The words “Blessed is the fruit of thy womb, Jesus” sounded slightly odd, polished from so much use, and recognizable only because of their context. So I stared harder. And there was the invariance. Every repetition was as identical as you like to the previous and the next. The practice of saying the rosary stabilizes the voice. (And every context has a context!)

So joint speech delivers another thing to think about. The Holy Grail of the phonetician was lying there in the quiet murmured prayers of an old woman, not in the laboratory attempt to capture letters in the voice. Over time, this observation grew into a recognition of the ubiquity of chanting prayers and mantras using beads. The Catholics have their delicate beads, small and smooth, in groups of 10 with isolated beads between each decade, and a small tail with some concluding prayers. But the Buddhists have their beads too. The mala, for that is the more common word, are recited in contemplative practices such as intense studying of the sutras. Buddhist beads are smooth, and come in a single loop of 108. Shaivite beads (Siva is my favourite Indian God, there I’ve said it!) are called Rudraksha Mala and they are large rough seeds, whose wrinkled skin is prized. Again there are 108. Beads do not do sectarianism. They are used by Vaishnavites, by Jains, by Moslems, by Orthodox and Roman Christians, each with their own symbolic embedding. We can actually paint a rough sketch of one world based on the beads. It is conventional, and usually misleading, to try to identify modernity and its complement, or East versus West, or first and third worlds, or developed and undeveloped countries. We seem only to perpetuate myths and create fictional others to whom we stand in opposition with such talk. But there are the bead users, and they are not the whole of humanity. Many people do not use beads. But a great number of people do, and we can indicate them using the figures of religion and culture. Within this world of bead users, the Jews stand out, precisely in not having the beads (though Jewish culture has no problem with elaborate ritual, chanting, and all the other stuff that comes with beads). I have found this

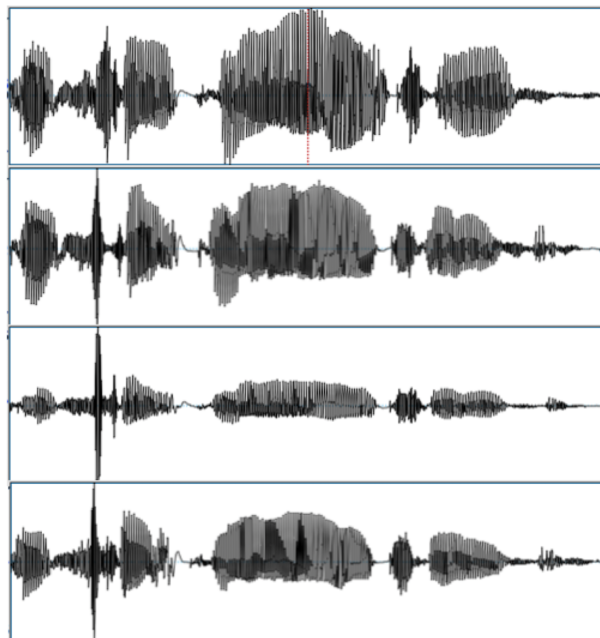


Figure 4: 'Blessed is the fruit of thy womb, Jesus" repeated four times, absolutely invariantly.

puzzling, and still do. Jews are definitely within the world of the bead users, and not outside it. Andaman Islanders are, presumably, outside it (I haven't checked), but the Jews are inside it. Thus this absence is striking.

9 The theoretical impasse: Language as a problem

I speak of an alienation from the discipline of linguistics. To be fair, phonetics is a place of enormous tension between the incarnate and the symbolic. If language can be said to exist, speech certainly grounds the symbolic, the disincarnate. It is understandable then why speech is examined for things we believe to be of the mind. The phoneme is central here, and a small digression might now be in place, before venturing into very uncertain territory.

The idea of a phoneme is the idea of a letter is the idea of a single sound that can be combined with others to produce language. Trained in the scientific tradition, I believed, and was taught, that this powerful idea had been best or originally fashioned by the Prague School of linguists from the mid-20th Century. I had learned the way that science had constructed the phoneme from a book by Jakobson, Fant and Halle, which systematically ordered sounds—with different treatment for vowels and consonants. This continued an undertaking that came from the 19th Century, to find some kind of culture-invariant structure underlying the sounds we make. This had given birth to the International Phonetic Alphabet, which aspired to order the world's consonants into a coherent structure, a table, that could make manifest all sound contrasts. Fig. 5 makes the scientific ideal clear: A finite set of elements are ordered along two (or more) dimensions, and the 'meaning' of each element, its role when it appears for real, is given by its place within this systematic array. This kind of thinking had made possible the prediction of an absent element, and

then its subsequent discovery, for example when the planet Neptune was found, or any one of the many elements predicted and then discovered or created (magic!) in the last hundred years. It is a very powerful way of thinking and ordering things, and it has been very influential in guiding our thought about the relation between speech and the symbolic order.

the international phonetic alphabet (2005)

consonants (pulmonic)

	LABIAL		CORONAL					DORSAL				RADICAL			LARYNGEAL
	Bilabial	Labio-dental	Dental	Alveolar	Palato-alveolar	Retroflex	Alveolo-palatal	Palatal	Velar	Uvular		Pharyngeal	Epi-glottal	Glottal	
Nasal	m	ɱ		n	ɳ		ɲ		ŋ	ɴ					
Plosive	p b			t d	ʈ ɖ		c ɟ		k ɡ	q ɢ			ʔ	ʔ̰	
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ			ħ	ʕ	ħ	ʕ̰
Approximant		ʋ		ɹ			j		ɰ						
Tap, flap															
Trill	ʙ			ʀ							R				
Lateral fricative				ɬ ɮ											
Lateral approximant				l			ʎ		ʟ						
Lateral flap				ɭ											

Where symbols appear in pairs, the one to the right represents a modally voiced consonant, except for murmured *f*.

Shaded areas denote articulations judged to be impossible. Light grey letters are unofficial extensions of the IPA.

1 H															2 He
3 Li	4 Be														10 Ne
11 Na	12 Mg														18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te
55 Cs	56 Ba	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn	87 Fr	88 Ra	89 Ac	90 Th	91 Pa	92 U
89 Fr	90 Ra	91 Ac	92 Th	93 Pa	94 U	95 Np	96 Pu	97 Am	98 Cm	99 Bk	100 Cf	101 Es	102 Fm	103 Md	104 No
107 Nh	108 Ta	109 Ta	110 Ta	111 Ta	112 Ta	113 Ta	114 Ta	115 Ta	116 Ta	117 Ta	118 Ta	119 Ta	120 Ta	121 Ta	122 Ta

- Known in antiquity
- also known when (Lavoisier published his list of elements) (1789)
- also known when (Mendeleev published his periodic table) (1869)
- also known when (Seaborg published his periodic table) (1945)
- also known when (IUPAC published its periodic table) (2000)
- also known when (IUPAC published its periodic table) (2012)

Figure 5: Top: International Phonetic Alphabet (2005 version). Bottom: Periodic Table of the Elements.

We can note in passing that vowels never quite succumbed to this systematic mode of ordering, because the tongue moves in a vast uncharted space without clear demarkation points. Consonants are describe with respect to occlusions that interrupt the airstream, and occlusions have places. The fluid tongue that shapes the air of utterance does not live in a box in an ordered table.

But this development of the idea of an elemental sound in speech arose in what we might call a modern scientific context. I did not know how long the idea had been around, and what importance it bore. The more I looked around at what speech had to offer, the more suspicious I became of the story I had imbibed in my training. My PhD supervisor, Bob Port, had finished his long and productive career and arrived at a similarly sceptical point. One of his final contributions was a keynote address with the evocative title *All is prosody: Phones and phonemes are the ghosts of letters*. Another boldly claimed that *Phonology is not psychological and speech processing is not linguistic*. I am not alone in my suspicion that something important needs to be illuminated here, and that phonetics is the point from which it can most clearly be seen.

So I wrote a little book about chanting, suggesting that this way of looking at things seemed to exceed the manner in which science, as I had learned it, looked at our utterances, our speech, or, as Sophie likes to call speech: our carefully controlled exhalation, for speech happens only on the outbreath. And I left the world of mainstream linguistics and my association with the main phonetics journal, which is called the *Journal of Phonetics*. I was at an impasse, when, once again, my friends provided the way forward.

Specifically, Sune Vork Steffensen, who works in Denmark, ran a journal that had one leg within the orthodoxy I was seeking to leave, and one outside. The journal *Language Sciences* specialised in misfits like me, who loved language and speech and culture and diversity and puzzles, but who wanted, as far as possible, to explore aspects of our language that found no recognition within the orthodoxy, or ideas which threatened it. I would like to give an example of two of the sorts of articles that the journal dealt with, one published, the other not published, that might give a sense of the spirit with which Sune led (and leads!) the journal, and what the landscape was that we found ourselves in. To really get a sense of that, I need to conjure up a strawman that I wrestle with every single day—a dream I am caught in, along with everyone else, that we recognize as a dream, but also as common-sensical and real, in some sense. I will describe the dream as the White Dream, and it presupposes English as a clear medium of communication, psychologically characterisable persons, number as the digits of calculation, the zero sign, and the dollar. That is rather a lot of things, and disbelieving in any one of them is relatively easy. For example, we can disbelieve in the dollar when we recognize that it is an invention, that there are situations where its seeming inexorableness can and must be dispensed with. (If you are dangling from a cliff and I pull you up, we do not, or should not, negotiate a price beforehand.) To disbelieve in all of these at the same time would be glorious and render you unfit for social participation and conversation. So let us simply call it the White Dream, and note that the field of Generative Linguistics, with its ideas of Universal Grammar and structure without meaning as the basis of language is one of the defining characteristics of the White Dream.

In the first article, Sara Petrollino⁴ undertook an anthropological task that is somewhat familiar. In its White Dream format, Western psychologists go out to far flung places in search of human universals. One of the favourite things to think about when we think about people, minds, and meaning, is the topic of colour. Colour has driven our best philosophers mad, and there is no accepted story one can rely on here. When the White Western Psychologist goes to far-flung shores of alterity, they bring with them tools for unlocking the architecture of this assumed shared common cognitive realm: they bring colour tiles, and they arrange the tiles in front of their ‘informants’ (who must think the whole game a bit odd) and ask them to name the colours. On a plastic tile, the notion of colour takes on a particular form, readily associated with a name: That is ‘blue’, that is ‘green’ etc. Based on the answers obtained, cultural variation is identified with slight differences in naming colours (that is, in employing colour names), while there is much that brings out our universal character. The salience of red, for example, seems clear to everybody, while blue and green are especially ill-bounded. (The White Dream notion of colour goes back to Newton and forward to the album cover of *The Dark Side of The Moon*: it is the prism that produces a rainbow. In such tasks, nobody asks about brown. The rainbow text gives much to think about.)

In the spirit of the journal, however, Sara Petrollino and colleagues went out to remote areas, not to find out how these people worked and why they were like the researchers themselves, but to be genuinely ignorant (ignorance is an epistemic virtue, not a sin) and capable of learning from the encounter. They did not bring colour tiles. Instead they took note of how the Hamar people of Southwest Ethiopia described things. Now, before we succumb to the dream, we must not only identify ‘people’, but the ‘real world’ of those people. The Hamar are cattle pastoralists. They live outdoors for the most part, and they live among their cows. The Hamar describe situations and patterns drawing from this world. There are no colour tiles here. Indeed, the Hamar have no world

⁴Petrollino, Sara. "The Hamar cattle model: the semantics of appearance in a pastoral linguaculture." *Language Sciences* 89 (2022): 101448.

for ‘colour’ (The authors note that there are many languages for which this is the case.) There are many terms for the job of describing how something appears, and the terms come, in the main, from the cows. Thus we find terms that the English speaking authors had to translate as ‘roan,’ ‘patched,’ ‘small spots,’ ‘forehead patch’ and so on. The familiar (Newtonian) dimensions of hue, saturation and value were not much use. Black, red and white had specific words, but for the most part, no transparent translation of the terms used was possible. This recognition is important: that others have words that are not translatable into English! Forgetting this is part of the White Dream.

The other paper that can illuminate the spirit of the journal *Sune* nurtures must remain anonymous. It was rejected because it was perfectly good linguistics, but it lay entirely within the White Dream. Let me lay out just a little history of the notion of Language and the (Christian) Bible. In the age of colonialism, the European and the European notion of dominion and control spread to vast areas. Everywhere the colonial forces intruded, Christian missionaries went along too, and they brought their precious book which we can call the Bible (recognizing that there is no single Bible). And they preached. And they translated. This business of bringing the Bible to unfortunate natives drove much of what was to become a science of translation. We recognize ‘languages’ now by looking up DuoLingo and seeing which languages are available. Most are also the name of political entities: Dutch, German, and so on. Everyone also knows that our speech is much more varied than this systematic view can acknowledge. The notion of language as a self contained system which will later underpin the whole of structural and generative linguistics relies on the notion of a ‘system.’ The boundaries of those systems were drawn at this time, and those boundaries were established when the work of translating the Bible had to be started afresh. Those entities we treat as real, the categories of DuoLingo, the distinct ‘languages’, these were created by this work. It is 2025, and as I write this, the world of field linguistics which provides basic linguistic data on all the world’s languages, is still run by a Christian missionary organisation, and though the overt reference to Biblical translation is hidden. This is how the notion of ‘language’ was constructed.

(A parenthesis on scripture is needed here: Moslems know that the Quran is the speech of God and is untranslatable. The Jewish canon is not yet closed. The Vedas are chanted. The Bible has a unique position here and its influence is everywhere in the White Dream that thinks translation is a matter of mapping from words to words, instead of the work of attempting to reexpress an original creative impulse. Translation, English as a transparent vehicle for communication, and the notion of ‘language-as-system’ are important components of the White Dream I wish to try to excuse myself from, and can’t.)

The second paper believed in translation to an extreme extent. It went so far as to claim that the authors had uncovered some basic words which were perfectly transparent, which could be found in every language, and which pointed to universal human cognition. We recall the embarrassment of the psychologists taking coloured tiles to far flung shores. The first, and clearest example of a word which was perfectly translatable across languages, and which thus pointed to a universal, was claimed to be ‘I.’ That was a position the authors had argued themselves into in prior work. They appealed to the journal to express an extension: they were now going to reveal some more perfectly translatable words. What could go wrong? The word “God,” they asserted, existed in all European languages and meant the same thing in all of them. They were now going to argue for a basic alphabet of human cognition that included atoms such as “mother,” “father,” “wife,” and “husband.” The journal did not accept the paper. I wrote a small paper called “Language as a Problem” and left the journal, reluctantly, but I could no longer be involved in this kind of

business. The dream is very hard to escape from. Sune saved me for a while, but I had to leave.

10 Four domains and Jerusalem

Joint speech was compelling, and there was no method for looking at it except to attend to context and more context and context within context. Luciana Longo was a student of mine and we laid plans to drop her into a Summer School on the West Coast of Ireland to be a fly on the wall and to observe all and any joint speech that occurred. Covid regulations and institutional sclerosis did not allow that to happen within the short period we had available, and so Luciana got her degree and moved on, learning Arabic in Jerusalem. But you can't stop the inquiry. She started dropping into schools in Israel and the occupied West Bank to observe joint speech. The technical term made this business sound somehow respectable, but of course she did not at all know what she was looking for. That turned out to be most useful. She persisted, and so I went and joined her and we discussed and stared at things together. Jerusalem is a great place for joint speech research. It has a little bit of history behind it, there is quite a bit of prayer and chant, and no shortage of protest. Luciana was keeping an eye on the schools, with the help of Zvi Bekerman of the Hebrew University of Jerusalem and he had been keeping an eye on the toxic effects of the White Dream there. Specifically, he had spent a career pointing out the dangers of psychological language in the education of children, language that imposes and insists on an identity, that must then creates others to whom one stands in opposition.⁵ The White Dream includes the elaboration of individualistic psychology where each person is forced to be distinct (we are not distinct), to have an identity (we do not have an identity, but we can build one and then go to war), and to construct an invariant and closed 'I.' Zvi helped us think things through.

These discussions led us to sharpen our toolbox. The definition of joint speech had changed a little by now. The original locution "multiple people saying the same thing at the same time" was not (and never will be) sufficient. We had long noted the erasure of any distinction between spoken words and musical notions in chant. Chant refuses any such distinction. Sometimes it has the prosody of speech (e.g. reciting a collective oath), sometimes it is clearly sung (The magnificence of Hildegard von Bingen's music) and frequently it is neither (Happy Birthday, U.S.A.). So we tried "multiple people uttering the same thing at the same time." That erased a remaining problematic tie to the White Dream of 'language.' This version doesn't quite extend to some situations I would like to include (the collective breathing during Sufi chanting is an important landmark for me at this stage), but we do what we can.

In our discussions, we noted that joint speech had the power to orient us towards activities by which persons construct shared worlds. There is no homogeneous domain from which to sample chants, but 20 years of observation had by now made it clear that wherever there was chanting going on there was something creative and vital to observe, and that we could identify four important domains that we were drawn to. In these domains, people were assembled with some kind of common purpose. They assembled, and when they assembled, they communed. Communion was a word that felt funny in my mouth when I first introduced it, but I now see that it is vital. "Communion is prior to communication" said the mathematician George Spencer-Brown. You can shout at an earthworm all day and communicate nothing, but an arched eyebrow raised by one member of a loving couple who have lived together for years can communicate volumes. If there is

⁵Bekerman, Zvi, and Michalinos Zembylas. *Psychologized language in education: Denaturalizing a regime of truth*. Springer, 2017.

any meaning to communication it does not lie in discrete letters or bits of information. But there I go again, dreaming I am out of the dream.

We must use words to communicate, so we gave four terms to the four domains we had identified. The insufficiency of the labels will be immediately apparent, and we use the labels only to point to the activity. The four domains are:

- **Ritual** This includes all practices of prayer, mantra recitation, oath swearing, and more. Nobody has ever been able to define ritual, but the approach taken by Roy Rappaport in his magnificent work *Ritual and Religion in the Making of Humanity* draws out many of the contours of this domain. Ritual is time-out-of-time. Rappaport describes the Cybernetics of the Holy. He discusses self-referentiality of ritual, semiotics, performance, Logos and science. The book is a wonder. Rappaport had also spotted joint speech: “the performance of more or less invariant sequences of formal acts and *utterances not entirely encoded by the performers* [which] logically entails the establishment of convention, the sealing of social contract, the construction of the integrated conventional orders we shall call Logoï” (Rappaport, 1999, p. 27, emphasis added). (Again, the question of authorship and utterance arises!) In an essay entitled “The obvious aspects of ritual” he notes: “It is, perhaps, important in this regard that the coordination of persons in ritual performances is often much tighter than usual . . . Indeed, the levels of coherence . . . often achieved in ritual is *more typical of organisms than of social groups*. It is important that unison is a common feature of ritual, and it may be made comprehensive by engaging several sense modalities and modes of expression simultaneously.” (Rappaport, 1974, p. 213, emphasis added). Let ‘ritual’ stand for the first domain. This is assembly with common purpose and orientation of one kind.
- **Protest** This domain is also assembly with common protest and orientation, but done in the streets, or wherever necessary, unprotected, vulnerable. Protest has its own forms, its technologies that are always developing, and, as Jeremy Woodruff showed, the chants in protest are rich and saturated with meaning that escapes any linguistic characterization.
- **Football** As if Jerusalem did not have enough to offer, my visit there coincided with the final of the World Cup. Football chant had not escaped our attention. Sepp Gumbrecht had written a wonderful book⁶ on the crowds in football, as a philosopher and a football fan, sensitive to the way collectives have been disparaged by philosophers and theorists, sensitive to the rawness of the struggle and the proximity to violence. Another philosopher and football fan is Steven Connor of Cambridge, who had already given me much to think about in his book *Dumbstruck: A Cultural History of Ventriloquism*, where the association between voices and singular human speaking subjects was troubled. (Who is speaking?) He is an Arsenal fan, and contributed the observation that Arsenal fans are coralled through an underpass on their way into the stadium, and the acoustic givens of the underpass inevitably cause the chanting to begin, there, every week. By football I mean what Americans call ‘soccer’, but chanting reveals something about sports itself: where the sport is team based, such that individuals can be substituted while the collective retains its identity (transfers!) and where teams enjoy cross-generational support (I support the team my father supported, etc.), we get chanting traditions. No such traditions arise in snooker or tennis, where the individual is the focus.

⁶Gumbrecht, Hans Ulrich. *Crowds: the stadium as a ritual of intensity*. Stanford University Press, 2021.

Steven wrote his own piece on ‘choralities’⁷ which rhymed terrifically with everything we had observed to date.

- **Primary education** Most of us will have learned our multiplication tables through chanting. This practice seems very widespread and venerable (though some chafe). This drew our attention, and we looked a little further. We stepped outside the White Dream to look at madrassas with similar naivety. Madrassas teach children to chant the Word of God, the Quran. This is not in the White Man’s dream. It is not English, or translatable. To chant the Quran from memory is a formidable achievement within Islam, and great, almost divine, attention is given to the niceties of the musical prosody of the chant. Understanding is different. In schools where Arabic is not the ambient language, the chanting of the Quran is taught on the basis of sound, not interpretation. In both cases, within and without the White Dream, sacred values were being transmitted through chanting. Maths is a sacred value of the White Dream.

Here are the four domains of communion and assembly, illustrated vaguely using the notion of a quasi-object (Michel Serres), which is not a thing, but more like Lebowski’s rug: the thing that holds the whole thing together. Note that protest does not have a protective enclosure, that there are oppsotional dynamics in sports chanting, and that the chant of the children is driven, not by their bodies, but by the values of their local society. Sune was kind enough to publish this article which hangs on such intangible (but obvious!) characteristics, and which makes its argument mainly using an odd picture. I live in hope that the twin terms ‘assembly’ and ‘communion’ may find service where ‘communication’ seems to fail us.

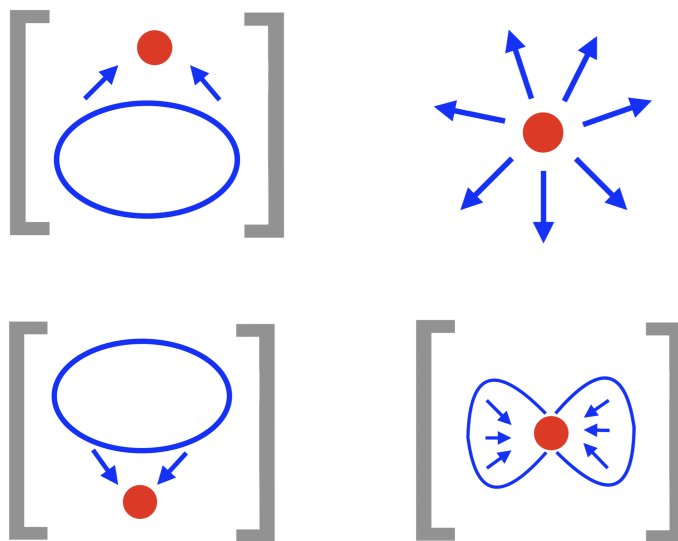


Figure 6: Four domains illustrated using quasi-objects: clockwise from top left: ritual, protest, education, sports.

⁷<https://stevenconnor.com/choralities.html>

11 Goodbye to rationality

Rationality is such a low goal. Reason is no substitute for reality. But phonetics continues, now, to enlarge. I no longer pretend to know anything at all. Phonetics gets involved in creation ‘myths.’

The Upanishads are fonts for Indian thinking, developed exegetically in the dharma dialectic between schools that a European eye can only see through a sectarian lens, picking out the divisive terms of Buddhist, Shaivite, Nath, Jain, etc. Those are all sectarian labels. What is at stake is nothing other than what science seeks also: Access to reality, and perhaps guidance in the conduct of one’s affairs, for a living embodied being. Here, I find my ignorance of the production of utterance magnified.

In the Taittirīya Upaniṣad, I read:

Om! We will now explain phonetics—phoneme, accent, quantity, strength, articulation, and connection. That describes the field of phonetics. (Chapter 1, Verse 2)

That’s it. That’s all we get here, but there is mystery lurking. The Maṇḍūkya Upaniṣad, which is very short, is devoted entirely to the cosmogenetic syllable Om. It begins “OM—this whole world is that syllable.”

Now we are very far from home, and from our specific dream. Perhaps this is another dream. Certainly, there is a lot of talk of phonetics, but not as I learned it. This dream has not lost the body as the White Man’s dream has. Utterance, speech, and breath pervade the Upaniṣads. There is the goddess Vac, speech herself. Speech, and chant, mix freely with gods, with Death. And in the midst of this unfamiliar territory the phonemes come back and join in. These are not the phonemes of Chomsky and Jakobson. But chant is our theme, not English, and this centrality of speech in many of the Indian schools (I use the term loosely) is far from capricious. It is developed and worked out in countless volumes, from the ancient grammarian Panini to the present day. The terms ‘phoneme’ and ‘letter’ are used interchangeably, and each is described with respect to place and manner of articulation, which is familiar territory.

Mantra lives here. The Gospel of Saint Matthew records an invective against “vain repetitions” which was picked up by Protestants to slander Catholics and their rosaries. Repetition seems inefficient if our uttering is understood as the business of communication or information. Even if she were seated on her throne in heaven, there can be no news in yet another repetition of the Hail Mary. Mantras have also been described in this way by outsiders: muttering over beads. Useless. Inefficient.⁸ But step inside the scholarly discourse about mantras (for scholarship and study are by no means confined within the White Dream or the domain of English), and we find goddesses, we find potent seeds that can be placed on the body during rituals that effect the transformation of the yogi’s body into a divine vehicle. Mantras, it is said, are transmitted within initiatory lineages. Funnily enough, as a young Irish boy I met this too. At age 14 I accompanied my mother to a short course on Transcendental Meditation, where we were taught to sit and observe thoughts silently. To aid us, we had a mantra whose mental repetition gave a place to come back to as the mind wandered. My mind wandered, but I completed the course and was given a small initiation ceremony, given simple dignity by a fresh flower, apple, and handkerchief, and I was given my own personalized mantra. Without fear, and confident that I damage nothing, I will ignore the

⁸The media saboteur Marshall McLuhan was an enthusiastic rosary reciter and toured applauding the virtues of the rosary.

injunction never to say the mantra aloud or to pass it to another, because if I say the mantra was *ima* I have told you nothing.

In looking around for expertise or concern with joint speech I had come across the practice of repetition of the Jesus Prayer among members of the Syrian Orthodox church (among others). Here, the short prayer “Lord Jesus, Son of God, have mercy on me” is repeated silently and constantly. It becomes a treasure that the worshipper has that they carry within them. It is of great importance to those within these denominations, and because we are naturally sectarian creatures, it is ignored by those without. But this pointed in the same direction. This was mantra too, in Orthodox form, curated, treasured. The catch-all category of ‘prayer’ includes such practices (and a great deal more) and has been entirely missed by investigation by the White Folk. There are foolish experiments in which groups pray for divine intercession and a control group does not pray for divine intercession, and the scientists can find no evidence of divine intercession, and so prayer is ignored, misunderstood, and rendered at best harmless. To treat the gods so is standard within the White Dream.

So I did not in the least mind abandoning scientific publication in English with all the attendant difficulties. There was more to be done, there is vastly more to be done, here, if we are careful. But first we must be suitably ignorant and open.

12 The plot thickens

Back to scripture! The Sefir Yezirah is a part of the (still open) Jewish canon. A vastly influential text of formidable obscurity, it has prompted exegesis and contemplation with discursive effects that have spread through Hermetic landscapes. We find its influence in Kaballah and in magic, throughout the discredited but lively pockets of esoteric learning that have always permeated the European realm and beyond. This text seems to reward a suitably imaginative hermeneutics.

And here comes phonetics and letters again. This time, the letters are not the Himalayan letters (We cannot, in this way of following things, distinguish between Tibetan and Sanskrit). They are the Hebrew ones, again an alphabet charged with magic and mystery. Chapter 4 of this enigmatic text confronts us with the following:

Twenty-two foundation letters: three mothers, seven doubles, and twelve simples. He hewed them in spirit, carved them in voice, fixed them in the mouth in five place: Alef, He, Het, Ayin [ahacha]; Bet, Vav, Mem, Feh [bumaf]; Gimel, Yod, Khaf, Qof [gikhaq]; Dalet, Tet, Lamed, Nun, Tav [datlanat]; Zayin, Samech, Tsadeh, Resh, Shin [zsats?ras]. Alef, He, Het, Ayin are pronounced at the end of the tongue at the place of swallowing; Bet, Vav, Mem, Feh between the teeth with the tip of the tongue; Gimel, Yod, Khaf, Qof on the palate; Dalet, Tet, Lamed, Nun, Tav on the middle of the tongue and pronounced with the voice; Zayin, Samech, Tsadeh, Resh, Shin between the teeth with the tongue at rest. (Chpt. 4, verse 3)

What is going on here? Once more, these letters are emphatically not the English letters. These sacred phonemes create reality within the text. This is a divine creation myth. Again, my weak training in phonetics did not train me for this, but phonetics draws us here, too. There are other creation myths. Indeed, who could confidently claim to know the many ways in which people of most varied being have contemplated their origins?

But we of the White Dream are familiar with the claims made of speech in such myths, even if we thought they were mere myth and thus not worth taking seriously. *Fiat lux!* Thus is creation uttered into being. In the beginning was the Word. Another uttered universe. This is where I began to realise that the Voice Onset Time, meticulously measured in phonetics laboratories from waveforms on screens, might signify something else.

The potency of the voice is foreign to the objectifying activities of the modern scientist, but the potency of the voice is familiar to every speaker. Its recurrence in creation myths around the world provides testament.

13 Taking stock

As a young would-be scientist, I dutifully measured voice onset time, and found it lacking in relevance (to me). How things have changed! I do not have words to describe the dream I both inhabit and rail against. I have called it the White Dream. Certainly, it is a background presumption of what others call ‘scientific modernity.’ It is probably also picked out by pointing roughly to the ‘Abrahamic traditions’ which cannot shake the notion of a soul, no matter how much they try to materialize and rationalize that as ‘mind,’ ‘cognitive system’ and the like. To a llama from Tibet, the Abrahamic imagination clings to a fatal kind of individuality that emphasises a discrete autonomous entity I refer to with the pronoun ‘I,’ presumed existant between birth and death. This pervasive assumption prevents exchange of wisdom between us. Tibetans (I use the term loosely) are excellent epistemologists. Buddhist epistemology is making small inroads into scientific work, in a field pointed to as 4E cognition, or embodied cognitive science. But the chasm is vast. I am more ignorant than when I started out, but richer too. Some things I suspect but cannot talk in public about without a great deal of clarificatory hedging include:

The notion of Language is at a great distance from the notion of utterance, the controlled exhalation we do together as we bring our common worlds into being. The valuable notion at the core of the term ‘language’ is better served within mathematics, and perhaps semiotics, for there we deal with distinctions that partition and structure universes. English is a problem, a technology, a Whiteness, that shields us from much we would better encounter and learn from. The llama I mentioned keenly observed that as one dives into Buddhist epistemology, one has to learn terms that are fundamentally untranslatable. Thus one finds oneself picking up a technical vocabulary that includes terms from Sanskrit, Pali, Tibetan, Chinese, and more. Thinking outside the White Dream is possible. If you traffic in cognition or consciousness, it seems to me to be required.

The utterance is a fundamentally creative act. To speak plainly, for the embodied being, the heart is the origin of time, the lifebreath begins with utterance, with the first cry of the new born, with all the chants, mantras, howls of our joint being. We simply are not well understood as individuals. Voice onset is the onset of everything. I could not know that as a young student. The voice does not have a single speaking subject. There is no answer to the perfectly rational White question “Who is speaking?”

And the things I have learned along the way find no place in the Great White Discourse, though they find plenty of resonance in conversation with others, because nobody actually lives entirely within a dream.