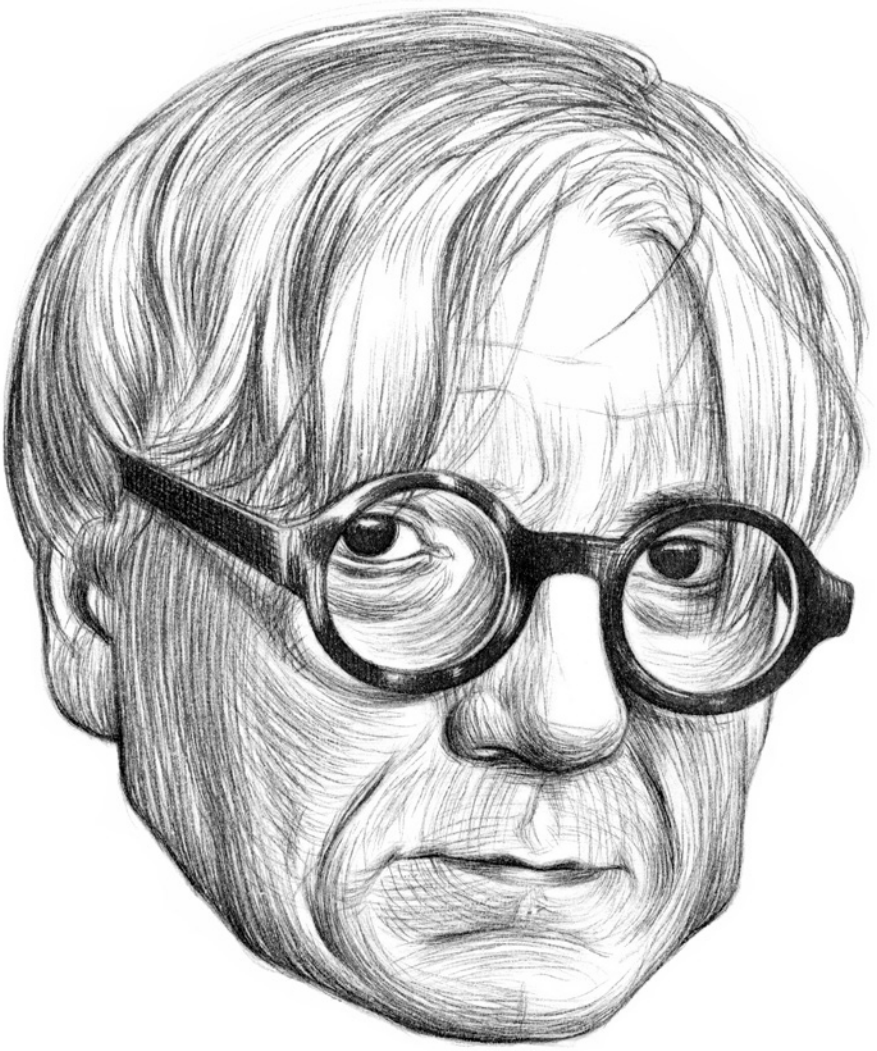


Cover and closing drawings by Simon Manfield

THE MONSTROSITY OF INFLUENCE

I remember a television documentary that BBC Channel 4 produced about Donald Judd after he died. I watched it in a video library in the basement of the monolithic concrete art block at Coventry University. Matthew Collings presented, looking very English, standing in the scorching Texan desert outside Judd's compound. A talking head contemporary of Judd's, Richard Serra, explained how Don began his career as an art critic and used his monotone reviews of painting exhibitions to foreshadow his own elementary sculptures. I can't actually recall, but I assume Collings delivered a comic parody of Judd's writing. Something like: There is a painting. It is red. There is another painting. It is green. A yellow band extends from its bottom left corner half way along its bottom edge … and so on. What I took from Serra's commentary was the idea that Judd's bluntly styled prose was a kind of hoax, preempting his appearance as an artist, setting the scene for artworks best described by their technical specifications. Serra, having introduced this notion, proclaimed it to be "very clever," in an admiring, slightly resentful way.

Another preemptive art-hoax is plotted in Scottish author Gilbert Adair's 1998 novel, *A Key of the Tower*. Sacha Liebermann is an academic, a leading authority on the work of the 17th-century painter Georges de La Tour, and he authors the definitive monograph on the artist. This apparently scholarly publication is in fact an elaborate deception whose entire, carefully disguised purpose is to canonize a La Tour painting that, so Liebermann reports, has been missing since 1885, and presumably destroyed. Secretly, in his studio, he has been forging — inventing — this work that in fact never existed. He and his accomplice, Beatrice, the wife of an unsuspecting art dealer, plan to "discover" the fictional painting, sell it to an unscrupulous collector for ten million pounds and elope. The hoax is foiled in the novel's conclusion, and Liebermann murderously undone.

Both narratives subversively anticipate artworks. Judging by Liebermann's fate in *A Key of the Tower*, Adair thought this contrivance distasteful. So consider the following inversion of the motif. There exists an obscure breed of writer known as the "anticipatory plagiarist." This term was apparently invented at the Ouvroir de littérature potentielle (workshop

for potential literature) to describe Charles Lutwidge Dodgson, the English author who, under the pen name Lewis Carroll, wrote *Alice's Adventures in Wonderland*. The Oulipo, a collective of writers and mathematicians founded in Paris in 1960, saw in Carroll an artist who had already formulated, a century before, the union of literary imagination and procedural writing devices that they aspired to. The label was a disarmingly mischievous acknowledgement of the derivative nature of THEIR project.

To be derivative is stigmatic, especially in literature. Gilbert Adair was the most exemplary of derivative writers, but, in the same mischievous spirit of the Oulipo, his relations with the canon were promiscuous and unfettered. His was a consistent artistic proposition: practically everything he wrote was a continuation of the work of another writer. He wrote novels, criticism and commentary (often on the subject of cinema), biography, translation, and adaptations for the screen. He wrote continuations of Carroll's *Alice* and J. M. Barrie's *Peter Pan*, recreating and extending their fictions and prose with extraordinary technical fidelity. He pastiched Agatha Christie in a trilogy of murder mysteries. He made a cultural translation of Roland Barthes' *Mythologies* (meaning he translated not the text, but the project, substituting Barthes' survey of French cultural references with English equivalents). He wrote a biography of the real life boy who inspired Thomas Mann's novella *Death in Venice*, and a novel, *Love and Death on Long Island*, whose portrait of obsession is an homage to Mann's book.

Continuity even pervades Adair's imagery. Returning to *A Key of the Tower*, the opening scene has two men — one of them the previously men unsuspecting art dealer, the other a lonely writer — driving in opposite directions along a remote road in the French countryside. It is night, and there is a wild storm. Lightning strikes the moment before they will pass, bringing down a tree and blocking the road between their two vehicles. At the roadside, the two strangers consider their predicament. Eventually, they resolve that the only way for them to continue their journeys is to swap cars. The exchange proceeds clumsily as the writer, having handed over the keys to his dilapidated Mini, struggles to master the controls of the art dealer's pristine Rolls-Royce.

Adair himself was consummately skilled in mastering the mechanisms of other writers. He achieved literary notoriety for his translation of Georges Perec's *La Disparition* from French into English, a project that occupied him for four years. Perec was the iconic member of the Oulipo, and his extraordinary novel is a lipogram on the letter "e," meaning that letter — the most commonly occurring in both French and English — does not appear once in its 300+ pages. Adair's version, titled *A Void*, is not only a translation of the original's narrative, but of its technical virtuosity. By translating a lipogram into a lipogram, Adair was outdoing the already impossibly prohibitive constraint of the original. I could go on, but the point is this: from entire fictions to a single letter, Adair worked in explicit relation to the writers he admired. He found a name for this practice via another French literary academic, Gérard Genette, who termed it "transtextual."

A FINISHED WORK REQUIRES URGENT RESUSCITATION

In September 2010, I initiated an interview with Adair. I wanted to understand his perspective on the canon, to situate him on a continuum between what I perceived as the two polar philosophies that his work might represent. At one pole, I imagined Adair a postmodern nihilist, scoffing at originality and engaged in a kind of morbid, joyless exercise in imitation. At the other pole, a more pragmatic position, in which the technical and aesthetic elements of a literary style are not bound to the identity of their author, but, once published, become vehicles for other writers to use in unforeseen ways. I got an email address from his agent and contacted Adair to propose the interview. I said that I would try to get the result published. He responded and agreed to a conversation over email.

James
Your first novel, *Alice Through the Needle's Eye* is written as a continuation of Lewis Carroll's two classics, *Alice's Adventures in Wonderland* and *Through the Looking Glass*. Like Carroll's originals, your text employs circular narratives and imagery. Your method seems to have something in common with the monographs of Gilles Deleuze, who described his portraits of philosophical predecessors Nietzsche, Kant, Spinoza and Bergson as forms of buggery: "I saw myself as taking an author

from behind and giving him a child that would be his own offspring, yet monstrous." Much of your subsequent fiction is informed by this idea of continuity. You refer to repetitions, continuing the journeys of others and so on. You have accounted for the history and status of this genre in your essay "On Transtextuality," but how did you proceed practically? What were the mechanics of appropriating Lewis Carroll's voice, for example? How did you begin?

Gilbert
I had never come across the Deleuze quote, but actually find it extremely interesting and pertinent, even if anatomically more than suspect (no one has ever had a child, however monstrous, by being taken from behind). I just hope it's not too urgent, as I'm a bit swamped with work at the moment and the question is not really one to which I can dash off an answer in ten minutes or so. Is there a deadline? Or are you able to wait until I have some proper breathing space?

I replied to say that there was no deadline, and — pleased that my reference appeared to have got his attention — that it would not be a problem for the response to be delayed. But nothing followed. I wrote a further polite reminder. No response. The interview had stalled. Until, a year later, I read Gilbert's obituary. I learned that in October 2010, shortly after our initial correspondence, he had suffered a debilitating stroke leaving him almost blind. A distinguished author blinded, unable to read or write without assistance: I recognized immediately the tragic parallel to the plot of Gilbert's own novel, *A Closed Book*.

Following this grave news, I reread "On Transtextuality." There, Gilbert casually outlines a project for a prospective apprentice to take up: "Consider Cocteau's celebrated monodrama … *The Human Voice* … whose middle-aged heroine makes one last, despairing telephone call to the lover who is on the point of abandoning her. It is possible to imagine a reworking of that play that would focus instead on the other end of the telephone connection, on the abandoning lover. Nor would it strain a competent writer's powers of invention to devise some new, transtextual title (*The Inhuman Voice*?). Then that writer has to get down to the workaday chore of writing the thing, from line to line, from page to page, and the problems he confronts are liable to be just as formidable as those

necessitated by a completely original scenario. The more so as he will naturally seek to avoid treading on the Master's toes."

It occurred to me that I might take *The Inhuman Voice* as an invitation to resuscitate my interview with Gilbert. First, I would need a method for simulating his voice. Anyone superficially familiar with Gilbert's fiction probably would have arrived first at the same obvious approach: to compose his responses to my questions by excerpting from his texts. Surely an insightful meta-narrative would be revealed that way. I assembled some promising passages and made a start:

James
Let me begin with the same question as before. You have accounted for the history and status of the transtextual, but how did you proceed practically? What were the mechanics of appropriating Lewis Carroll's voice, for example? How did you begin?

Found-Gilbert
Were one to see a horse with a man's head, one would not cry, "Look! A horse with a man's head!" but "Look! A centaur!" And were one to see a woman with a fish's tail, one would not cry, "Look! A woman with a fish's tail!" but "Look! A mermaid!" A centaur is a centaur, a mermaid is a mermaid. Mythological (in the word's more orthodox usage) as they are, such creatures have come to possess in our eyes (or in our imagination) their own compact and specific identities: if they are "singular," then it is in the sense both of uniqueness and oneness.

Was that Gilbert's head on Carroll's body, or the other way around? I looked for further clues in the texts and uncovered nothing. In any case, Found-Gilbert's enigmatic response had pointed out the flimsy premise in my approach. I had ignored an important distinction: Gilbert's writing is always more reconstituted than found, more *monstrous.*

Simulating Gilbert would require a more subtle appraisal of his work. In other essays I noted Gilbert rehearsing positions that he might have taken had our conversation continued. In "Black and White in Color" he criticizes the practice of retrospectively colorizing black and white films: "Since black and white cinematography is coterminous with a fairly precise

era of film history, colorization represents an anachronism as indefensible as would be the engrafting of a belated soundtrack onto a silent classic." "On Updating" is an equally severe dismissal of the early 2000s trend for productions of stage classics in contemporary scenarios — Shakespeare with mobile phones and the like. "No one has ever reorchestrated a Bruckner symphony in the hope of appealing to fans of Sondheim or Take That. No one has ever repainted a Rembrandt self-portrait on the grounds that the sight of a wrinkled old codger in a 17th-century smock would prove alienating to a spectator of the nineties." This last excerpt may contain the essence of Gilbert's sensibility. The friction contained in a sentence comparing Take That AND Bruckner was what seemed to delight him. Not Bruckner BY Take That, just simply Bruckner AND Take That AND …

This was an instructive realization. A recreation of Gilbert's voice must express something previously unsaid, however derivative. Accordingly, I devised an alternative format for the interview. I would write to friends and colleagues of Gilbert, asking them to respond to one of my original questions, as if in Gilbert's voice. I initially sent two emails, one to a noted Oulipian whose address I obtained through the Parisian gallery where he exhibits sculptures. No responses. Discouraged, I revisited Cocteau and *The Human Voice*. Perhaps I ought to take the speculative outline of *The Inhuman Voice* even more literally.

THE INHUMAN VOICE

CereProc is a digital voice synthesis production studio based at Edinburgh University. I discovered their work through an online video of a TED talk titled "Remaking my Voice," given by journalist and screenwriter Roger Ebert in 2011. Ebert had lost his speech while suffering from throat cancer. In the video, he demonstrates various voice synthesis systems and has family and colleagues speak for him, reading from scripts. He describes how his dissatisfaction with generic computerized voices led him to seek a more personal tool to express himself, and ultimately to the production of a unique simulation of his voice, made by CereProc, using archive recordings provided by Ebert himself.

Intrigued by the prospect of a posthumous digital synthesis of Gilbert's voice, I searched for suitable material. Gilbert's friend David Thompson kindly agreed to send me a copy of a compilation of clips of Gilbert on television and film that he had prepared for a memorial. Gilbert made cameos in film adaptations of his novels and appeared regularly as a critic and presenter on the BBC's *The Late Show*. In the first clip on the DVD, he is seen talking about Jean Cocteau again, standing in front of a Cocteau fresco, praising the various output of the French visual artist, writer, filmmaker, and boxing trainer. Expecting to hear a Scottish accent, I was surprised that Gilbert's delivery has a measured, southern English character to it. (I later learned that his speaking style is a good example of Received Pronunciation, the archetypal form of broadcast English.) He slips effortlessly between English and French as he recites titles and quotations from Cocteau. How fitting that his voice should conform to a conventional idea of neutrality. Nonetheless, Gilbert's diction did have some distinctive characteristics. He seemed to relish constructing certain words in his mouth, such as "evol-yew-shun" and "otto-matic."

I made contact with the executors of Gilbert's estate, again solicited by his agent. They responded to say that they did indeed have some audio, and asked for what purposes I intended to use it. I wrote enthusiastically, explaining that I was researching the possibility of making a synthesis of Gilbert's voice so that I could finish a conversation that I had started with him. This may have been poorly judged. No reply.

I also found references online to recorded lectures by Gilbert. An inquiry with one academic institution prompted an unsettling episode. On the morning of 22 October 2013, two years after his death, I received an email from Gilbert Adair. Subject: "My voice." It reads: "I got a message from Surrey University saying you wanted to get a recording from them because you're making a computer synthesized version of my voice. I'll probably be happy to give permission for this, but I am intrigued … could you fill me in a bit more on the project?" There is another Gilbert Adair, a poet. I suppressed my embarrassment and considered whether this was opportune. Should I ask this Living-Gilbert to answer my questions? Too farcical. Then should I ask Living-Gilbert to impersonate Dead-Gilbert for a recording? Have him try to learn and reproduce the character of Dead-Gilbert's voice? It does sound like something out of

a Nabokov novel. Dead-Gilbert might have appreciated that. I responded apologetically, explained my mistake, and did not hear from Living-Gilbert again.

Finally, with only the memorial compilation, I contacted CereProc and began an email correspondence with a software developer there, Graham Leary. As an introduction to CereProc, Graham told me about one of his projects that involved making a synthesized voice for Jason Bradbury, a television presenter known for his work on *The Gadget Show*. The voice was used in a robotic model of Bradbury's head and torso. Graham sent me a link to a YouTube video in which Robot-Jase, wearing a bow-tie, is introduced as the first robot to host a live UK television broadcast. Standing next to his double, Jason says: "He's the star of the show. I'm like a bit-part. FOR MYSELF. It's very postmodern." Graham disclaimed that the lip-syncing is not very good, and that CereProc did not work on that.

Graham was exactly what I had been looking for: an expert in voice synthesis with a mischievous sense of humor. Without hesitation (or announcement), I cast him as Gilbert in my improvised production of *An Inhuman Voice*.

James
I've been wanting to ask, what were the mechanics of appropriating ~~Lewis Carroll's~~ Roger Ebert's voice, for example? How did you begin?

Graham-Gilbert
I normally build voices for living subjects, who read from a script in controlled studio conditions. I need around 50 hours of recorded audio of a subject speaking to make a functional voice synthesis. However, as you know, in 2009 CereProc did make a voice for Roger Ebert, after he lost his speech. That voice was built entirely from pre-recorded material. We cut each of the recordings into very small fragments that are called "diphones." A diphone is the transition from one phone to another. So, if you take a word like "cat," the phones are "cuh," "ah," "tuh," and the diphones are "silence-to-cuh," "cuh-to-ah," "ah-to-tuh," and "tuh-to-silence"…

Speaking of cats, in *Alice Through the Needle's Eye*, Gilbert's Alice meets Ping and Pang, two Siamese who are joined at the tail and speak as one being, habitually finishing each other's sentences. "You took the very words out of my mouth!"

... We cut the recordings into diphones, rather than cutting them into individual phones, because there's less variability in the middle of a phone than there is on a phone boundary. The boundaries are where the most characteristic parts of each phone are found. These cuttings —some as short as one tenth of a second— are then stored in a database. Our software rapidly recombines the cuttings to render the speech. A voice synthesis is programmable in a text-to-speech framework, where text input into any supported software can be spoken aloud by the computer.

James
Are there other methods for building a voice synthesis?

Graham-Gilbert
There are many. The one I described above is a CONCATENATIVE voice synthesis. The actual sound is preserved from the original recording of the subject speaking. The synthesis is in the division and recombination of the sounds. One alternative, a FORMANT synthesis, builds an entirely virtual system by modelling the characteristics of a subject's voice and reproducing them digitally.

This distinction between Concatenative and Formant synthesis reminded me of the Deleuze quote I had shared with Gilbert. A Concatenative voice synthesis would reconstitute Gilbert's own words —taken right out of his mouth— yet *monstrously.* Puppet-Gilbert.

Something else Graham said resonated. In the production of a Concatenative voice synthesis for a living subject, there is an automated process for dividing the recording into its constituent diphones. Speech recognition software processes the recording, matching the audio to the script read by the subject. The computer can only process the audio if it recognizes the exact sequence of utterances found in the script.

Ping
If it knows what the speaker …

Pang
… will say.

This process is called "forced alignment." The computer is matching what it hears to the basic structure of the script, forcing the recorded sounds into the predetermined rhythm of the text. In the case of a deceased subject — since a dead person cannot read a script — the "script" must be produced after the recording. A transcript of a recording would be given to the computer so that it can perform its forced alignment. These contrary flows, *script → speech* alignment and *speech → script* alignment circled my attention back to preemptive hoaxes and anticipatory plagiarism. Gilbert would have made the connection, too. In *Love and Death on Long Island*, his disturbed protagonist writes: "My novel flowed as fluently from my pen as though the complete narrative had somehow been miniaturized in advance and injected into the nib and it were merely a matter of posing the pen over the paper, teasing each word, like a droplet of ink, off its tip and having it spill onto the blank page."

Another point of interest on the matter of transcription. There are established orthographies for the representation of phones and diphones, and the stresses on particular sounds in any given articulation of a word. Graham provided examples of phonetic orthography from CereProc's Received Pronunciation lexicon:

Input
k_a1_t

Output
Cat.

Input
k_a1_t_@0_n_ai0_n_t_ei1_l_z

Output
Cat-o'-nine-tails.

Input
k_a2_t_@0_k_l_i1_z_m_i0_k

Output
Cataclysmic.

The numbers indicate the stress: primary (1), secondary (2) and unstressed (0).

Finally, my conversation with Graham arrived at the decisive question. The memorial compilation is 38 minutes long, but unfortunately includes only around eight minutes of Gilbert speaking that are clearly audible, free from background music or ambient sound.

James
It is technically possible to make a Concatenative voice synthesis from only a tiny sample of source material?

Graham-Gilbert
Your material would probably only have partial diphone coverage. Confronted with a database that lacks certain diphones, there are successive levels of "back-away" that our software would resort to. In the ideal situation, the database would contain all of the possible diphones, and the software would recombine them in the largest available chunks. The software would try to preserve the recorded clusters of sounds as far as possible. If you typed a word like "caution," and the word "cautious" had been recorded in the script, the software would render the "causha" from the recording and then graft an "n" onto the end from another cutting. It doesn't always concatenate at the diphone level. If the database were missing some particularly obscure diphone — such as sh_ch, zh_ch, or zh_zh — the result might only very occasionally manifest in the rendered speech. The software would substitute the missing diphone with another similar sounding diphone from the same database. In case of a very patchy database, once potential substitutions were exhausted, the software would be programmed to contact an alternative database, producing an audible slippage in the rendered speech as the software inserts diphones from the secondary source. It would work, but it would sound very odd, like an aural version of a ransom note.

A Robot-Gilbert produced using only these eight minutes of material would be able to speak only a severely limited English, bound by the diphones Gilbert happened to utter in the clips compiled by David Thompson.

GILBERT AGAIN

Gilbert wrote that during the four years that he worked on translating Perec's lipogrammatic prose, his perspective on the constraint changed. "At the outset, the language I had to contend with seemed to me frankly 'crippled,' as though, in losing all its 'e's, it had lost several seriously major limbs. Then, as I started to feel more comfortable with its warped angularities, I became inclined to regard it less as 'crippled' than as 'disabled,' diminished by its loss of limbs, to be sure, but still capable of living a rich and full life."

Depending on the accent of the speaker, spoken English contains 40 to 44 phones. Received Pronunciation English has 44. The number of diphones equals the number of phones squared. There are also diphones between each phone and silence, making the total number 45×45 (-1 as silence_silence is not required) = 2024 diphones in Received Pronunciation English. The number of missing diphones in Robot-Gilbert's Concatenative voice synthesis is around a third of these, including some common diphones. Robot-Gilbert's language is DECIDEDLY crippled. Could even Gilbert have mastered such a constraint? Robot-Gilbert is not programmed to use a substitute database in the case of missing diphones, so whenever a diphone he cannot utter or substitute appears in text, he simply skips over it, without any allowance in timing. As I familiarize myself with his controls, it seems I might have to abandon any idea of writing a sophisticated monologue using all of Robot-Gilbert's diphones. I resort to trial and error, experimenting with the transformative effects of Robot-Gilbert's strange concatenations on text.

Earlier in the process, I had come across an auspicious fragment of dialog that I hoped would make the perfect conclusion to my exchange with Found-Gilbert. Gilbert co-wrote Portuguese director Raul Ruiz's 1981 film *The Territory*, in which a group of tourists embark on a hiking trip in a vast forest. They employ a guide called Gilbert. Guide-Gilbert's

motivations are unclear, but during the first day of their expedition, the group begin to doubt him. "He's leading us in circles," one of the group protests, as they pass the same few landmarks repeatedly. That night at their campsite an argument ensues, and Guide-Gilbert abandons the group. Days later they are still lost, their efforts to find a way out of the forest increasingly desperate.

"Did you walk in a straight line?"

"I tried to, but it's impossible."

"Maybe we should try walking in circles if we want to go in a straight line."

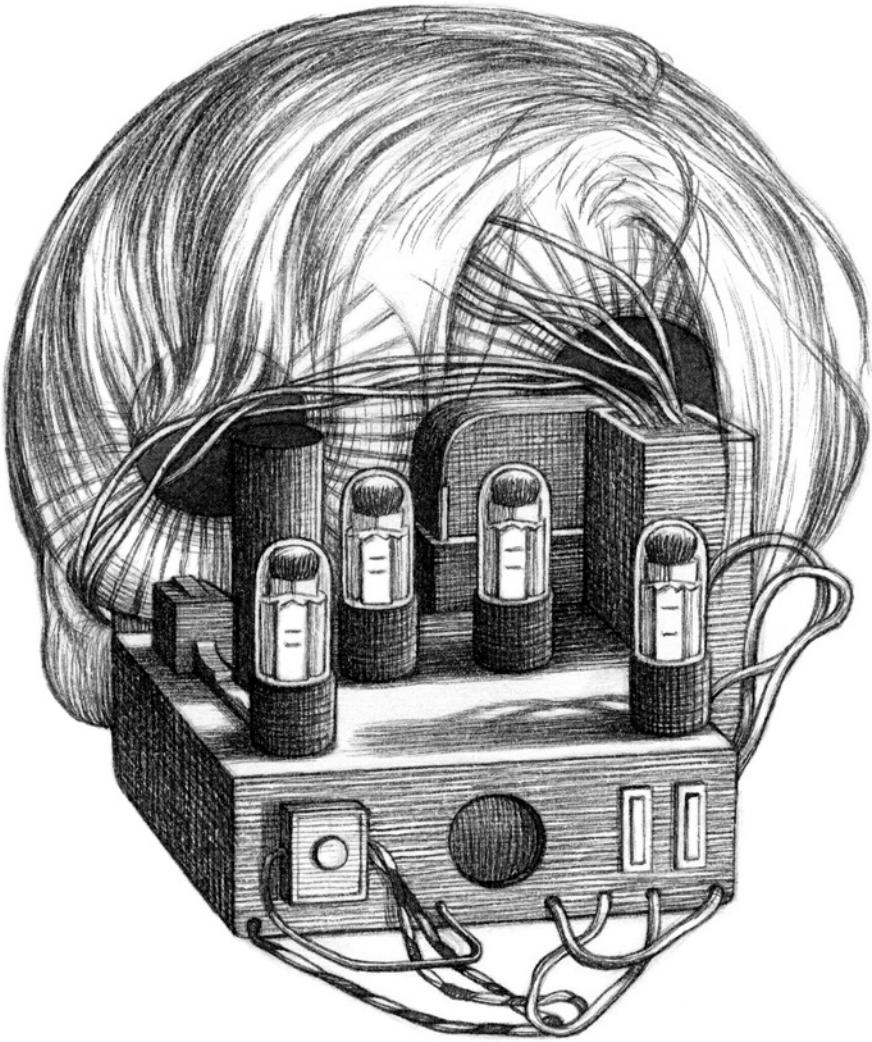
Delirious, they stumble upon Guide-Gilbert's body, his death unexplained. The discovery forces a terrible decision. Cannibalism might be their only option for survival. Objectifying the body helps the group to rationalize their actions: "Why do we still call him Gilbert? He's just a corpse now." The meat is cooked and passed around. As the adults eat with grim faces, two boys playfully argue over the provenance of the meat. One says to the other: "This is Gilbert, not your Gilbert." The other replies: "I've got a better piece than you. It's more Gilbert than your Gilbert."

Let me introduce Robot-Gilbert with a demonstration. I type this last line of dialog into TextEdit, in phonetic code. Missing the diphone b_@0, Robot-Gilbert can't resist treading on his Master's toes:

James
i1_t_s m_oo1 g_i1_l_b_@0_t dh_@0_n y_oo1 g_i1_l_b_@0_t

Robot-Gilbert
It's more Gilbot than your Gilbot.

*



BoTSL#7 2014 JUL 28 7:30 PM