

THE TALE OF 'NAMAJIMAKU', or Why The Dog Slept on a Banana

by Simon Prentis (published in Issue 26 of the J-Net Bulletin)

I was in Boston when the idea first came up, a month before the Japan leg of Paul McCartney's ongoing world tour. As his sometime "official interpreter" and cultural trouble-shooter (I'd accompanied Paul & Linda to Japan on their 1990 and 1993 tours, and kept in touch through Japan-related business ever since) I'd been asked to come out and help with some advance media promotion for Japan. A pretty cushy number, all things considered; but my smug grin soon evaporated when the tour director sidled over to me in a quiet moment and said, "By the way, we want to have Paul's words translated and displayed live on the video screens behind him when he speaks on stage in Japan. Sort it for us, would you?"

Sort it? Do you have any idea what you're asking?

They didn't, of course. As is so often the way in celebrity la-la land, someone comes up with a big idea, and with no-one able or willing to say no, it mutates within microseconds into *Something That Must Happen*, with very little thought as to, well, how it might actually be done. Not that the original impulse was bad – better communication with the audience in a land noted for subdued responses could only be a good thing – it was just that few in the entourage had enough experience with languages (including their own, in many cases) to understand why it might not be all that straightforward. After all, it wasn't as if there was going to be a script; and although I knew that McCartney tends to come up with the same general patter between songs, he'd made it clear he didn't want to have his style cramped by having to say the same thing every night – so the Covent Garden surtitle trick was not going to be an option. Simultaneous interpreting we could do; but simultaneous texting as well?

I politely explained, as best as I could, that though this particular pig might be made to grow wings, it wouldn't necessarily fly. Not only had nobody succeeded in getting it off the ground before, there were all the usual reasons why they were never likely to. Even if the various technical issues could be overcome, there would still be a significant delay between the spoken word and any subsequent display on screen.

But it soon became clear I was wasting my breath. As far as Sir Paul was concerned, I was given to understand, “can’t” is the little girl who lives on “won’t” street. Never mind the theory – until I could demonstrate to everyone’s satisfaction that it was absolutely and definitively impossible, I was expected to comply with the royal decree. And woe betide me if I couldn’t, because none of the inner circle were queuing up to be the messenger.

It’s not that they didn’t want to be helpful. Within hours of my initial reaction, minions had been despatched to check with both the US Chamber of Commerce and the British Embassy in Tokyo, who confirmed that if it could be made to work it would indeed be a first, whilst seconding my reservations. But this did nothing to curb the enthusiasm for the project, and I was subjected to a never-ending stream of well-meaning suggestions: “Don’t they do that kind of thing every day at the UN?” “Why can’t you just listen to the English and type it out in Japanese?” “They have court stenographers in Japan, don’t they? Just get a bilingual one and have them do it..” “What about close-captioning? There’s a girl in the crew who lived in Japan once who says they do it all the time...” “If it’s so difficult to type Japanese then why don’t we get him to speak slowly and they can translate each word as he says it...” Etcetera, and, er, quite.

So despite being employed to give advice, it was clear that it was not going to be accepted unless I had chapter and verse to back it up. I started making enquiries. I first of all checked with the simultaneous agencies, and unsurprisingly drew a blank. Next stop was the court stenographer’s association. Even with my relatively limited experience of simultaneous interpreting, I’d done enough to know that the mere thought of adding typing in Japanese to the workload would be a recipe for cerebral meltdown, but it turned out that in any case – something I didn’t know – stenographers use a specially modified typewriter that produces shorthand, not characters, which is typed up later for the record. Further enquiries with secretarial agencies confirmed that the fastest speed for typing Japanese using a word-processor is around 80 characters a minute, about one-third to one-quarter of the speed at which people actually speak. And as for voice recognition, don’t get me started...

The only suggestion I didn't initially follow up on was close-captioning technology, as I was pretty sure it wasn't used in Japan (despite the widespread use of on-screen text in recent years, it's largely for effect and all done in post-production). Nor had it been encouraging to learn from the BBC, when I called to find out how it was done in English, that they prefer real-time audio typing because it is faster and more accurate than voice recognition (which in any case has to be done using a trained operator who shadows the broadcast output.) But in the interests of protecting my exposed rear-end, I thought I'd better just check. I called up a producer friend in NTV, who, after initially confirming my suspicions, agreed to make a few enquiries. The next day he got back in touch to say he'd heard that both NHK and NTV were rumoured to be running a pilot live close-captioning scheme for the benefit of "the audially challenged", although it wasn't yet generally available. It was worth a try.

Real-time audio-typing in Japanese has apparently been something of a holy grail ever since word-processing was first invented (I remember being amazed to discover when I first arrived in Japan that it took hours to for the school secretary to type out a single page of Japanese using a device that looked like a cross between a miniature printing press and a delinquent pantograph). But it turns out that painstaking research carried out in a basement lab in Waseda University over the last 20 years has now not only cracked it, but found a commercial application through the entrepreneurial skills of a company that specialises in supplying live texting (albeit at enormous cost) to NHK. And cost being no object at this point, it was to them I turned.

When all the necessary introductions had been made and I eventually made my way to the nerve-centre of the operation, a nondescript 3rd floor office opposite NHK, it was hard to believe that the solution had taken so much time. Because although the devil (and the corporate secret) is in the detail, essentially all the system requires is a team of six people typing in sequence to produce a real-time output. In very simple terms, what happens is that three teams of two people each are assigned to a particular project, each pair consisting of an audio typist who types in phonetic Japanese and an editor who converts the incoming phonetic data to characters as necessary, these three teams working together to leapfrog down a sentence in bite-size phrases. So for example, once team A begins editing, team B starts typing the continuation of the sentence, and then when they start editing team C kicks in, the

intervals (which are pretty fast, in practice) being timed by shouting out key syllables so that when team C start editing, team A are ready to take over again and so on down the line.

Between them they produce a stream of raw text data that lags behind the audio input by about 2 seconds. This is sent down a phone line to the TV station where it is smoothed out (NHK buffer the input it and display it line by line, like subtitles, whereas NTV prefer to deliver a character-by-character display like an old telex) to overcome the jerkiness with which the text is actually output by the typing team. Unfortunately, this process introduces a further delay of around 2 seconds – not really a problem for an audience unable to hear the audio component of the broadcast, but crucial in a concert situation as we would have to factor in an extra delay for the interpreting. My initial prediction, when I was still naively hoping we might just be able to do something by teaming up a simultaneous interpreter and a superhuman audio typist, had been that we would be looking at a speech-to-screen delay of around 3-5 seconds, allowing a 2-3 second lag for the interpreting and 1-2 seconds for the typing. That may not sound like much, but if you're waiting for an audience reaction, it could be a very lonely moment. Which was why my advice had been that I didn't think it was really worth trying; but when you already have everything else money can buy, the chance to do something that's never been done is apparently irresistible. And Macca was determined to go for it.

So we set out to see if we could make it work. I decided that to minimise the delay we'd have to dispense with the smoothing software and deliver raw data to the screen, jerky or not. But there wasn't much time to fine-tune the system, especially since we needed to have a trial run before we arrived in Japan; given the constraints on getting into the Tokyo Dome for the first show (the whole stage with its giant video screens and state-of-the-art lighting and audio system had to be set up from scratch in less than 24 hours, a minor miracle in itself) there would be no time to run a test in Japan. And so ten days before we were due to go live in Japan I found myself standing on stage in Los Angeles after the band's regular pre-show sound-check, attempting to head-scratch Sir Paul through the complexities of the system so he could understand why it wasn't happening quite as simultaneously as he had hoped. After all, he had the microphone in his hand, and could see the characters coming up on the screen; why should he care that what was actually happening was that his

words were being beamed via satellite to a tiny room in Tokyo where a nervous interpreter who'd been kept up waiting all night was translating them down the phone to a roomful of frenzied typists whose text output was then bounced back again across the Pacific to Los Angeles?

"Is it working then, Si?" was what he really wanted to know. "I mean, for all I know it could be saying "Who does this guy think he is...?" Well, I couldn't deny there were teething problems – apart from anything else the scrolling wasn't working quite right – but he'd got the hang of the delay pretty quickly, and was learning to play with it, so it was good enough to be worth a try. And when it came to it, I didn't want to be the one to tell him no either, any more than anyone else. Besides, we had a few days to run more tests and fix the display issues. I gave it the thumbs up.

Of course, when it came to the first night, things didn't go so well. First of all we'd had to make a last-minute adjustment to the font size so that the text would be visible even from the 'nose-bleed' seats (the Dome holds about 50,000, and McCartney had only been playing to 20,000 seater-venues in America) and for some reason this made the scrolling problem even worse. Not only that, the flow of text was just too jerky, meaning that in some cases it scrolled straight off the screen before anyone had time to read it. And on top of that, McCartney had decided to have some fun. Determined to demonstrate what his shiny new system could do, he announced to the audience that he was going to depart from his normal routine. So to the surprise of us all, he suddenly started throwing in things like "My dog slipped on a banana. My cat has green hair..." making up ever more bizarre images as he went. It didn't help that the interpreter heard his initial offering as "My dog *slept* on a banana" (which is how it came out on screen) but in any case his culturally opaque bon mots were met with a deafening silence. You could almost hear the calculation going on in the audience's heads: "Did he really just say that? If so, why? Could it be a joke? Should we laugh? What if the translation is wrong? If it wasn't a joke and we laugh, that will be embarrassing. If it was a mistake and we laugh, that will be embarrassing too. No one else is laughing. Better not laugh, then...")

"Hey guys! Anyone awake out there?" The nasal edge of McCartney's native Scouse slashed through the silence as he attempted to defuse the awkwardness. "Oh well, never mind. Let's get on with the show..." It was an embarrassing moment, but

fortunately the audience still loved the concert, getting up on their feet and remaining there for the whole 3 hours. So in the band bus on the way back to the hotel he was in a conciliatory mood, and I didn't get the rough ride I was expecting. Fortunately we had a free day before the next show, and a chance to put the system right. I hurriedly organised someone to rewrite the faulty data transfer protocol, and after working through the night we had it cracked. This, together with adding a extra line to the text display so that there were three lines for the text to scroll over instead of two, more or less fixed it. The only other thing was to get the typists to make their individual chunks even smaller, to try and reduce the jerkiness of the flow.

And hallelujah, it worked. Apart from the odd early cut-off due to a video screen switching over-ride beyond my control, from then on in we had a fairly seamless display of what became known to the team as 'nama-jimaku'. And it was quite a thrill, despite myself, to sit out in the audience each night thereafter for the remaining shows and hear the collective intake of breath once the penny dropped that this actually was live, not pre-programmed surtitling. The delay was still there, of course. It was hard to measure precisely, since about a third of the audience were getting what he was saying directly in English anyway, and the reaction times overlapped, but I reckoned it was averaging about 3 seconds. That's still long enough, but the boss was pleased – and that was the main thing. And in fact it was probably quite helpful for people who weren't quite sure they'd understood the English it to have the meaning confirmed for them after they'd had a moment to think about it. And for those who didn't get the English in the first place, it was definitely added value.

So the little pig did fly to market, after all. I'd originally joked to the tour director that if we succeeded in making it work I wanted to have a scene added to the 'pre-show' theatrics (a 15-minute 'Cirque-du-Soleil' type performance with actors in surreal costumes walking through the audience) with a winged pig chased by a man holding a pointed stick, but in truth one of the advantages of being Paul McCartney is that you can get the impossible done. Because unless he had insisted, it would never have happened. And although it may never happen again (the costs involved probably far outweigh the benefits for anything but a show on this scale) it was nice to know that for however brief a moment I'd helped to create a little piece of history – the world's very first Japanese nama-jimaku.