



Eric Vatikiotis-Bateson and the Birth of AVSP

Eight years after Barbara Dodd and Ruth Campbell's Hearing by Eye (1987)[1], "the first edited book on the psychology of lip-reading"[2], the NATO Advanced Studies Institute (ASI) Workshop, Speechreading by Man and Machine[3] was held at the Chateau de Bonas, Castera-Verduzan, France. This two-week meeting was the first interdisciplinary meeting devoted to the subject of speechreading. It was a one of a kind. And it sowed the seeds for AVSP research to come. Forty-five researchers (from twelve countries), whose work covered topics such as human perception and cognition, linguistics, neuroscience, computer animation, machine learning and computer vision, held forth in a never-to-be-repeated cavalcade of science show and tell. What follows is first an episodic account of scenes from the meeting, that begins with the author's initial encounter with Eric, from which the story shifts to the formation of the Audio-Visual Speech Association (AVISA), the establishment of the AVSP conferences, and the role that Eric played therein, and finally, it traces a few outlines of the bigger picture of Eric's impact on AVSP.

On a sunny Sunday afternoon on August 17, 1995, I waited at Toulouse airport at what I thought to be the pick-up point. Across the way I spied a man in baggy shorts and Hawaiian shirt in animated conversation with a Japanese colleague.

I looked around to see if I was at the correct place for the pick-up; when I looked back, baggy shorts and his Japanese companion were gone. It turns out they were at the correct pick-up point for the Chateau de Bonas, Castera-Verduzan, and I was not.



When I finally reached Bonas about midnight, there was baggy shorts sitting by the pool, beer in one hand, roll-your-own ciggy in the other.

“I hear you missed the bus. Are all Australians as stoop-id as you?”

I had now met, as formally as I ever would, Eric Vatikiotis-Bateson – researcher, thinker, agitator, shit-stirrer, and lovable loudmouth. I joined him for a beer, and for more than 20 years of research advances, ideas, and irreverent fun.

Bonas was an academic and social blast; 45 auditory-visual speech processing researchers at the top of their game from all walks (see Appendix 1): the charming can-do Benoît (Benny); the



erudite and productive Bernstein; Helena, the goddess of speech perception without speech cues; the young first-timer, PhD student Iain (“Wow, are all conferences like this?”) now working at Oculus; Piero, the Italian Talking Head with a Feeling Heart; the Stately Stork, looking down from on high for the critical visual speech features with which to augment auditory ASR algorithms (“is the F-tuck the answer?”); the Latin Midday Magnet (“Oscar, you have five women at table today!” cried Eric and Benny in unison); HMM Silsbee, whose daily dress code was predictable based on priors (he left his

baggage unattended at Toulouse and it went the way of the Rainbow Warrior – boom!); the incisive, quietly-(Queen’s English)-spoken Brooksy; Ruth, dancing slowly with Benny to the village fiddler, squinting against the smoke from the fag jammed seductively in the side of her mouth; Cohen, the hippy-length-haired ballroom dancer who built the hairless Baldi; the other Eric, the US-based Big Gun Eric the Viking, flown in on a quick raid to give us the industry view; Jordi the whimsical and model integrationist; The Green McGurk-Meister and meister of much more; the Masterclass Massaro, first and most logical (or was it fuzzy?) questioner in every talk, and King of the Mountain over the Alps ahead of the young Turk jobseekers; and Eric the Irrepressible, wandering around the conference room, hardly ever still, asking questions – not

to tell people he was there, but to make a damn good point –, then, when he felt like it, re-arranging the program so he could give an extra presentation.

Cast of named characters

Eric: Eric Vatikiotis-Bateson
 Benny: Christian Benoît
 Bernstein: Lynne Bernstein
 Helena: Helena Saldaña
 Iain: Iain Matthews
 Piero: Piero Cosi
 Stork: David Stork
 Oscar: Oscar García
 Silsbee: Peter Silsbee
 Brooksy: Michael N. Brooke
 Ruth: Ruth Campbell
 Cohen: Michael Cohen
 Eric (Viking): Eric Petajan
 Jordi: Jordi Robert-Ribes
 Green: Kerry Green
 Massaro: Dominic Massaro
 Narrator: Denis Burnham

Amid the fresh-picked figs for breakfast; wine carafes and two-hour lunches; tastings in the Bonas cellar avec our chatelaine, the Countess; the petanque knockout competition (did Dom win that, too?); the Bateson-Benoit comradery; the new friendships and post-conference trips to Saint Michel... amid, and maybe because of all that, Bonas was two weeks of damn good research talks (including the impromptu one) and resultant advances in the auditory-visual Zeitgeist. Forty-five scientists presented, discussed, compared, and critically evaluated their work on speechreading, bimodal speech perception by normal and hearing impaired listeners, and audio-visual speech recognition (AVSR) by machines from multiple perspectives – psychological, neurophysiological, phonetic, signal and image processing, biometrics, computer vision, sensory fusion, and didactic.

Chateau du Bonas, 1995, was the birthplace of modern auditory-visual speech research; of AVISA; and of the AVSP conferences.

The next year, 1996, Christian Benoît and Lynne Bernstein organised a special session (The Multiple Senses of Speech Perception) at the *International Conference on Spoken Language Processing* (ICSLP) in Philadelphia.

Then in September 1997, the first *Audio-Visual Speech Processing* (AVSP) workshop was organised by Benoît and Ruth Campbell – a two-day satellite of the 5th *Eurospeech* in Rhodes with over 100 participants from 18 countries.

Throughout, Eric and his sparring partner, Benny, were the perfect auditory-visual speech processing team – together they created great, oft-times fantastic and bizarre research ideas, leading to implementations in code, hardware or experimentation. But in April 1998 Christian Benoît died at the tender age of 42. We had lost a dear friend, and none of us felt this more deeply than Eric.

Later in 1998 Eric rallied and together with Philip Rubin set up the *Talking Head Website*, first proposed at the 1997 meeting in Rhodes; and with Jordi Robert-Ribes and Denis Burnham organised the 2nd AVSP at Terrigal, near Sydney, Australia.

At AVSP '98 Eric led the charge and on December 5, 1998, AVISA (the Auditory-Visual Speech Association) conceived earlier by Benny and Eric, became the 2nd ISCA Special Interest Group (SIG). Eric always shunned grandiose titles, so he pushed Philip Rubin into the Interim President position, with Eric himself and Denis as interim steering committee members.

The AVISA objectives that Benny and Eric devised are to link AV researchers; facilitate better understanding of processes of auditory-visual speech perception and production and the modelling of speech gestures for realistic animation; share knowledge of AVSP in as many languages as possible; and disseminate AVSP research knowledge worldwide. Almost 20 years later it continues to do just that.

Throughout his career Eric skirted around the walls of conference halls chewing on ideas and interjecting with sage comments; he set up collaborations and instigated or augmented new labs in Japan, Brazil, Canada, Australia, France, USA, Germany and more; he bred a generation of auditory-visual speech production and perception researchers; and he cajoled, criticised, abused and hugged, and drank and smoked, with his collaborators. In 2005, Eric hosted a memorable AVSP on Vancouver Island, with the topics of the invited lectures perfectly illustrating Eric's approach to what needs to be understood: A lecture on the structure and function of the jaw; one on modelling face and tongue biomechanics, and one to remind us what it's all about – 'Appreciating face-to-face dialogue'.

Eric played a critical role in establishing the AVSP community – AVISA and the AVSP conferences; they would not be here without him. Eric is gone, but he is larger than life.... We thank him, and those of us who knew him, love him still.

Denis Burnham, with input from
Laurie Fais, Ruth Campbell, Kevin Munhall, and Phillip Rubin and Editor-in-Chief, Chris Davis

References

1. Dodd, B. and Campbell, R. (Eds) (1987) *Hearing by Eye: The Psychology of Lip-Reading*. Lawrence Erlbaum Associates, London.
2. Campbell, R., Dodd, B., & Burnham, D. (Eds) (1998) *Hearing by Eye II: Advances in the Psychology of Speechreading and Auditory-visual Speech*. Hove, UK: Psychology Press, P. ix.
3. Stork, D. G., & Hennecke, M. E. (Eds.). (1996). *Speechreading by humans and machines: models, systems, and applications* (Vol. 150). Springer Science & Business Media.

Appendix 1 (Bonus: Chateau de Bonas attendees)



1 P. Kricos	2 K. Pichora-Fuller	3 P. Smeele	4 O. Garcia	5 J. Beskow	6 P. Cosi
7 M. Sams	8 D.G. Stork	9 D. Massaro	10 C. Benoit	11 J. Movellan	12 A. Adjoudani
13 I. Matthews	14M. Vogt	15 J. Luetttin	16 K. Green	17 S. Hiki	18 N.M. Brooke
19 C. Abry	20 P. Bertelson	21M. Cohen	22 J. Robert-Ribes	23 T. Coianiz	24 E. Auer, Jr.
25M. Piquemal	26 D. Burnham	27 P. Silsbee	28 M.-A. Cathiard	29 L. Bernstein	30 H. Saldaña
31 A. Fuster-Duran	32 R. Campbell	33 M. Hirayama	34 E. Vatikiotis-Bateson	35 M.E. Hennecke	

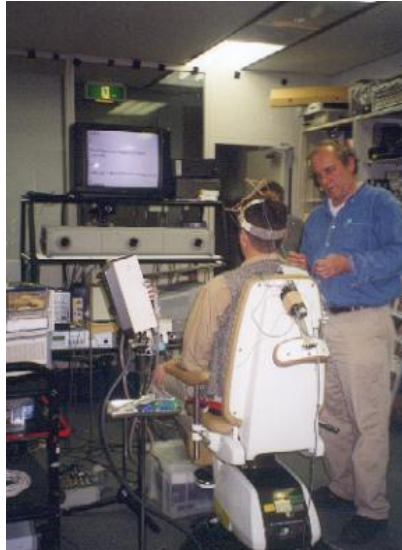
Not shown: Martine Cornuejols, Manuel Grana, Fabio Lavagetto, Si Wei Lu, Eric Petajan, Leandro Rodriguez-Linares, Robert Sokol

Appendix 2 (or was that the larynx?)

1.susPECT

mDB-A042susPECT.aif

Eric eliciting stress contrasts from an unsuspecting Australian with a hook electrode in his cricothyroid muscle



2.SUSpect
(with Eric comment)

mDB-A013SUSpect.aif