

# The Multimodal Melody of the Text: A Model of the Text Generator

**By Michael Mair**

*Ophthalmologist, Timaru Eye Clinic, New Zealand; Member, Work Group One (WG1) ISOTC/215 (the Modeling Work Group of the International Standards Organization Health Records Standards Committee).*

## **Abstract:**

This paper approaches the process of Multimodal text generation from two different perspectives:

- Brain processes which may be involved in the 'text' of transmodal' spontaneous discourse.
- A generic model for a prospective global Electronic Health Record standard (the ISO TC/215 committee).

The author conducted a micro-analysis of spontaneous dyadic interaction, obtaining a 'score' for Fundamental Frequency (Fo), phonetic analysis, head movement, and conventional script with a temporal detail of one tenth of a second. This data demonstrates many sound and movement clusters which are the property of the dyad, such as supra-individual rhythm, synchronies, melodic interweaving, and laughing. The trajectories of the movement shapes are most economically described as trans modal.

I attempt to model the production of this script out of known brain processes and their timings, including eye movement, head movement, and voice movement. I discuss how the continuously updating model of 'the present' that the interact ants share is itself supra individual and trans modal. I discuss fundamental frequency (prosody) as an artificial abstraction from the total kinesic trajectory of 'speech with movement', and account for the many 'functions' ascribed to it from this viewpoint. I suggest that there is a systematic relationship between melody and outcome in dyadic interaction, and that melody 'hard wires' the brains of interact ants in a communicative synctium.

I discuss some of the brain structures implicated in consciousness, and develop a 'brain in action' model of the process of text generation. I discuss 'mirror neurons' discovered in the cortex in this context, and the relation between core and cortex brain structures in the 'text generator' model. I challenge the dichotomy of 'emotion' and 'communication' and try to show how the 'text' of spontaneous human interaction is the trajectory between stable states, and that the 'affect' is a property of the models immanent in the interact ional stream.

I relate the limit to the number of variables involved in the logic in use' of everyday interaction (e.g. Miller's 'Law of Seven Plus or Minus Two') to the

dimensionality of space/time, and that of plans that act on it. I discuss the relation between the human 'virtual worlds' of sign systems and the more space/time bound animal worlds. The 'ontology' of the core brain functions is recognized as the core process of practical human interaction, as it is in other brain processes such as vision. The logic used by the 'text generator' is characterized as 'projective' on 4 D space/time.

I conclude by describing how this simple model of text generation has been deployed in the design of an electronic health record keeping system. I justify this design by the 'appeal to nature' described in the first part of the paper, and suggest that it works because it is in accord with the physiological process of text generation. I develop this concept further with reference to the 'model of models' General Domain Model' under discussion for adoption by the ISO TC/215 Health Records Committee, and to other international standards work in the Electronic Health Record area.

This timely convergence of theoretical approaches from different directions and for different purposes has far reaching implications.

Mike Mair  
August 5<sup>th</sup> 2001