

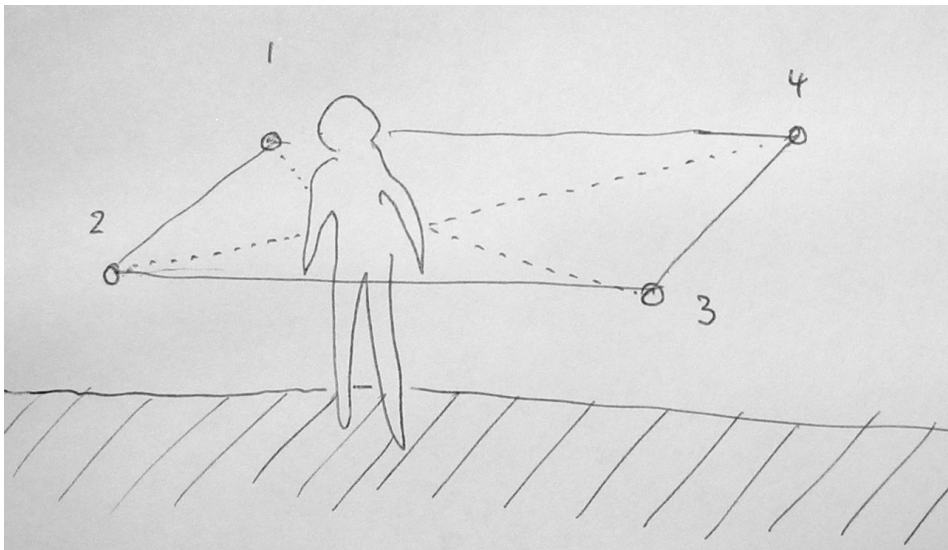
## Poly-sensing environment

### Summary of ideas, presented on 08/21/02 (2)

- The speech recognition consists of four different states that can be distinguished by the system and that have to be taken into consideration when designing a visual output:

silence	ambient noise	noise	trained patterns
no sound	low volume volume under threshold not analyzed	volume above threshold i.e. is being analyzed but matches none of the trained patterns	volume above threshold is analyzed and matches one of the trained patterns

- It is not only possible to train the system on words (i.e. "speech" recognition) but - after proper training - it should be also able to recognize musical patterns, environmental sound patterns, etc. (given that these patterns are fairly distinguishable).
- The activity of each node can be mapped over time and visualized in a timeline with several tracks, where each track represents a node (cf. graphical representation of activities in chat rooms on the internet in "chat circles", Prof. Judith Donath, M.I.T. Media Lab, see references)
- tracking a person's position using "audio tracking": cf. also Prof. Hiroshi Ishii, M.I.T. Media Lab and the project "Ping Pong Plus" (see references).  
the position of a person can be calculated based on either the time difference among the audio signals that the field nodes receive or based on the amplitude differences at each of the nodes.



In this simple sketch a person's x/y position can be calculated based on time/amplitude differences. It might already work with three field nodes in a plane (triangle). The more field nodes the more complex the tracking - even x/y/z tracking in 3d space might be possible with a few more field nodes.