

# Developing a Multi-Modal Corpus:

## Data coding issues

# Background Information-

---

- DReSS: Digital Records for eSocial Science
  - General aims and objectives
  - Specific linguistic aims and objectives
  
- HeadTalk:
  - Exploring gestural and verbal signals of active listenership-minimal responses in conversation.

# Requirements-

---

- To develop our understanding of head-nods and other gestures we need to consider:
  - 1) Data sources and collection methods
  - 2) Detecting and defining head nods
  - 3) Coding schemes and methods
  - 4) Re(presenting) data

# 1) Data sources and collection methods-

---



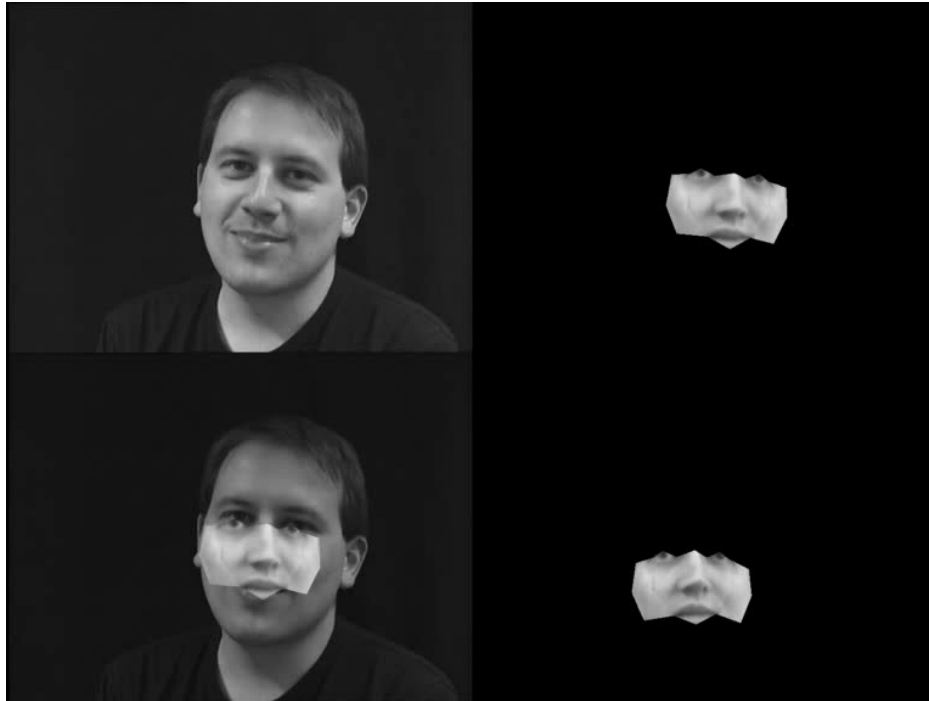
## 1b) Data sources... cont.

---

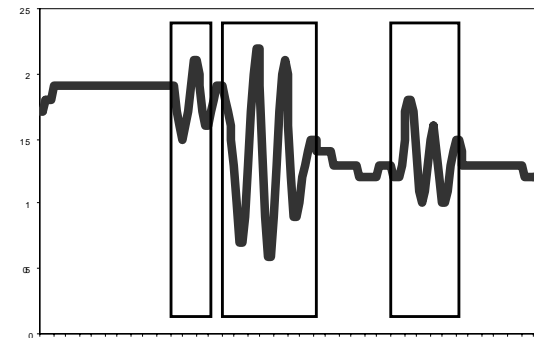
- Further data sets will include:
  - Further supervision sessions
  - TV interviews, PM Question Time
  - Map Tests and map tasks (run but others on DReSS)
  - Research meetings (for example DReSS meetings)
  - Lectures and conferences
  - Other gestures: laughter; handtalk; proxemics

## 2) Detecting and defining head nods

---



The model is fitted to each frame in the video, giving the 3D position and orientation of the head in each frame



Analysing the parameters of the model lets us find nods as oscillations of the head pose

### 3) Coding Data

---

- Coding schemes:
  - Schemes exist to handle verbal elements of interaction, but few integrate verbal with non-verbal
- Following extensive research on the multi-million word CANCODE corpus we have extended our backchannel coding scheme to include 4 basic functions.
  - Continuers
  - Convergence Tokens
  - Information Receipt Tokens
  - Engaged Response Tokens
    - » (O’Keeffe & Adolphs forthcoming)
  - These need to be developed for multi-modal linguistic corpora

## 3b) Coding.... a starting point

---



## 4) Representation & usability issues (1)

The screenshot displays the Replaytool application window. The main area shows a video player with two camera views of a man and a woman. The interface includes several panels:

- Annotations:** A list of time-stamped annotations such as "14:09:52.504 <\$F> HEADNOD".
- File selection:** A table listing files and their properties.
- headtalk.txt:** A text window showing a transcript of the video content with time markers.
- Play controls:** A control bar with a progress slider, a "Restart" button, and standard play/pause/stop buttons.

File	Mime type	Frozen	Start time
headnod1.mpg	video/mpeg	No	05-Oct-2005 14:09:46
headtalk.txt	text/plain	No	05-Oct-2005 14:09:46

00:00:00.000	Start of video
00:00:01.000	<\$M> and like you know like we said last time what I did was I looked at something like McCarthy's
00:00:06.504	<\$F> HEADNOD
00:00:12.218	<\$F> Yeah.
00:00:12.220	<\$M> words +
00:00:12.928	<\$F> yeah.
00:00:13.100	<\$M> as compared to speech +
00:00:13.362	<\$F> HEADNOD
00:00:14.500	<\$F> mmm.
00:00:14.958	<\$F> HEADNOD (half nod)
00:00:15.000	<\$M> and i cross referenced that with a book by Leech and a few others and that it's literally just
00:00:24.000	<\$F> uh-huh.
00:00:24.100	<\$M> and so I just looked at the McCarthy list looked at the Leech list and um <M> looked picked out
00:00:24.198	<\$F> HEADNOD
00:00:34.530	<\$F> HEADNOD

## 4b) Representation & usability issues

---

<\$F> **yeah.**

<\$F> **\*NOD\***

<\$M> or two American newspapers and looked at the er pre= premodifying adjectives +

<\$F> **uh-huh.**

<\$M> of any differen= of any word related to a a nationality or a or +

<\$F> **okay.**

<\$F> **\*NOD\***

<\$M> a country.

<\$F> **oh that's interesting** +

<\$M> and that what they were doing was basically saying well let's look at all these countries +

<\$F> **mmm.**

<\$M> or the word nation or the word

<\$F> **\*NOD\***

<\$M> state things +

<\$F> **mmm.**

<\$M> like that an= and when they say premodify they did it in very loose terms just looked at any adjective.

<\$F> **mmm.**

<\$F> **\*NOD\***

<\$M> immediately to the left on th=

<\$F> **right okay.**

---

# CONCLUSION