

Analysing spoken corpora:

Methodological issues and technical challenges

Communication:

- Beyond language: Communication as a 'complex network' of 'semiotic channels' (Brown, 1986: 409)
- These channels are multimodal
- There are many possible, 'different, independent, pragmatic and semantic functions' of signs making them specific to their (Argyle, 1975)

- Type

- Function

- Context of use

- Effective communication relies upon the receiver successfully detecting, processing and understanding these interactive 'signs' in its given context of use.

Research foundations:

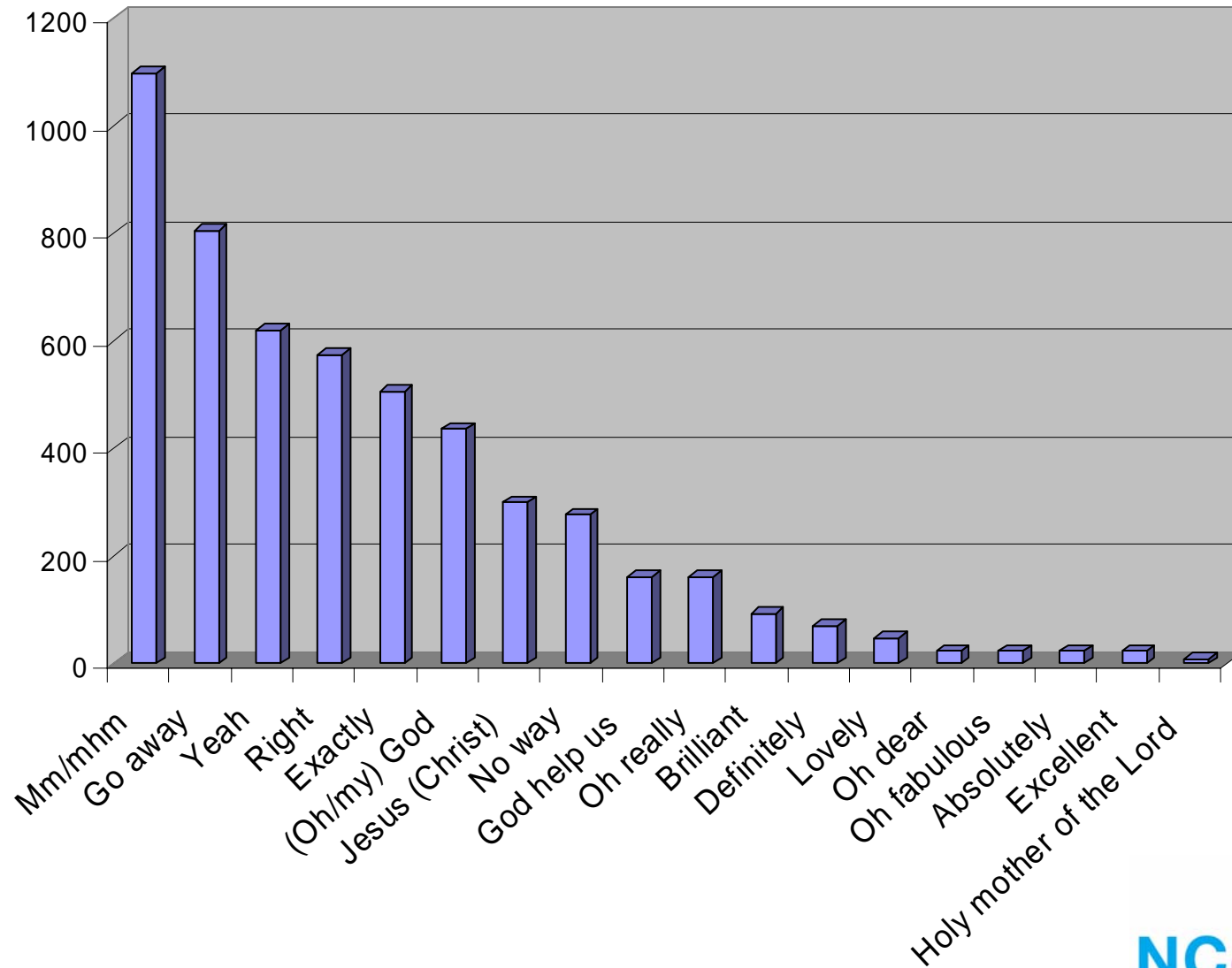
- Non-verbal, multi-modal behaviour plays an integral part in determining the meaning & function of linguistic components & units.
- The challenge of understanding the nature and role of gesture-in-talk (or in context as it were) is a substantial and *very* long term one that is of central concern to language studies (Kendon, 1994).
- Current methodologies in language data analysis need to be extended to include an integrated exploration of both verbal & non-verbal patterns of social interaction.
- To this end video corpora recorded in specific contexts have to be developed to illustrate the relationship between verbal and non-verbal behaviour, as well as the technological requirements for an integrated way of annotating different aspects of communicative events.

Possible sites for a multi-modal approach:

Analysing listenership

- A backchannel item is a short response token that does not take over a speaker turn and is not a response to a question.
- Backchannels are realised through a finite set of linguistic forms which can be extracted from corpora.
- Backchannelling also includes non-verbal response tokens, non-vocalised kinesic signals, and proxemic movement as a means for hearers to register and evaluate what is being said.

Backchannel forms in 40,000 words of conversation



Backchannel functions in discourse:

1) The Continuer (maintain the flow of the conversation)

S1: The way this was done was a Scottish lady who lived across the road from us.

S2: Yeah.

S1: And she would soak some grey wool. A length of grey wool in a saucer with olive oil.

S2: Yeah.

2) Engaged Response Token (conveys an emotive response)

S1: And do you wax under your eye?

S2: Yeah under both eyes I get waxed.

S1: Holy mother of the Lord.

S2: I get waxed under both eyes and my chin.

Outstanding issues:

- How does body movement relate to language forms used to signal active listenership?
- How does body movement relate to functions of backchannels?
- How does intonation relate to body movement, forms and functions of backchannel behaviour?
- How does body movement relate to different social contexts and groups of speakers/listeners?

Aims:

- * To record multiple modes of communication in a natural context.
- * To record both the individual & synchronised patterns of speech / head movements simultaneously, within the same frame of reference.
- * Recordings to be accurate and able to be replayed & annotated in the future.

Approach to research & Corpus development:

To develop our understanding of head-nods and other gestures we need to consider:

- Data sources and collection methods
- Detect, define, encode head nods
- Re(presenting) data

Data sources:

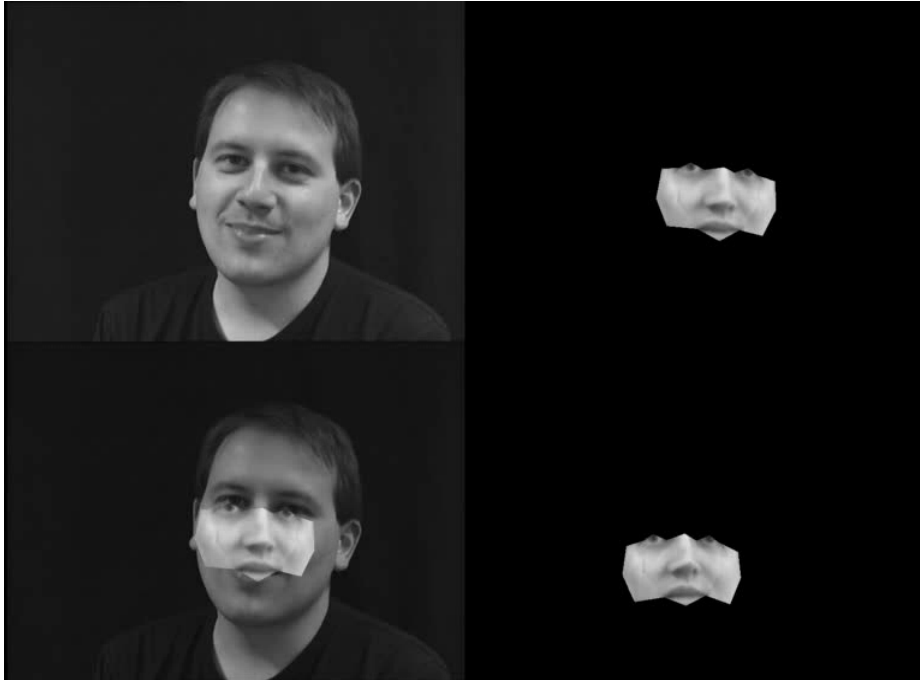
Data sets include:

- supervision sessions
- TV interviews, PM Question Time
- Research meetings
- Lectures and conferences

Other gestures will include:

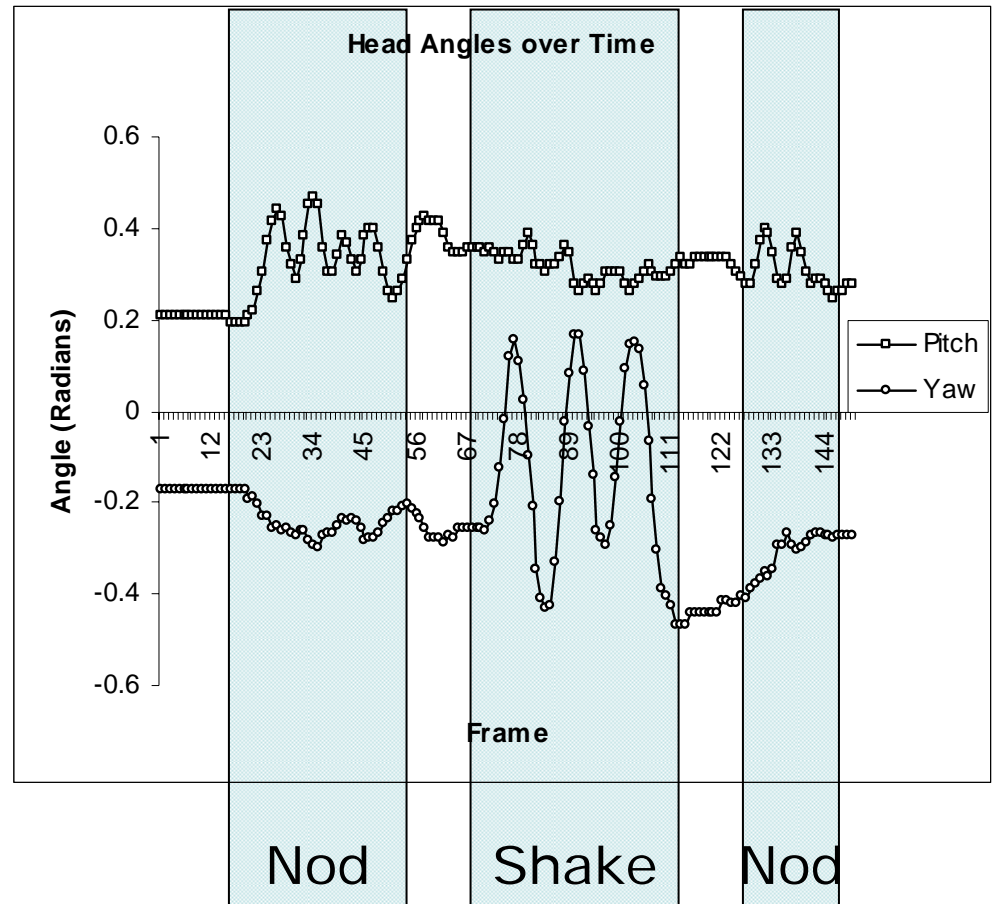
- laughter; handtalk; proxemics

Detecting head nods:



- 3D morphable model of human head built from image data
- Model is matched to each frame of an image sequence to track the head movements
- Slow, but potentially very powerful

Detecting head nods continued:



Encoding head nods:



Coding schemes and methods:

- Coding schemes supporting the identification, representation and analysis of different elements, components and units that exist in spoken discourse proliferate.
- There is a lack of such schemes for marking-up non-verbal elements or for integrating it with verbal elements, those that exist focus on computational tools.
- Current schemes are generally developed to match a specific research need or are aimed at a particular user group.

(Re)presenting Data- the issues:

- Current tools tend to focus either on the **management** of data or upon the **processes** of coding and annotating previously collected data.
- There does not appear to be a tool available to support the **integration** of these individual processes, supporting the research process from recording, through organisation of records, coding and analysis.

(Re)presenting Data:

The screenshot displays the Replaytool application window. The main area shows a video player with four quadrants of a video recording. The 'Annotations' panel lists time-coded events. The 'File selection' panel shows a table of files. The 'headtalk.txt' panel shows a transcript of the video content.

Annotations

- 14:09:52.504 <\$F> HEADNOD
- 14:09:59.362 <\$F> HEADNOD
- 14:10:00.958 <\$F> HEADNOD (half nod)
- 14:10:10.198 <\$F> HEADNOD
- 14:10:20.530 <\$F> HEADNOD
- 14:10:32.905 <\$F> HEADNOD
- 14:10:42.346 <\$F> HEADNOD
- 14:10:48.488 <\$F> HEADNOD

File selection

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

headtalk.txt

```
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
00:00:35.500 <$F> uh-huh
```

Some future challenges

- Re-classification of categories in the light of new visual and audio evidence?
- Sharing of tools, coding schemes and recording mechanisms with other research communities
- Ethics and data