

TalkBank and PhonBank

Brian MacWhinney

CMU - Psychology, Modern Languages, LTI
SDU - IFKI

Goals of this talk

1. Introduce TalkBank
2. Explain the theory of meshed time frames.
3. Analyses to evaluate this account.
4. Explain how the TalkBank principles derive from this vision.

CHILDES and TalkBank

	CHILDES	TalkBank
Age	24 years	8 years
Words	44 million	8 + 55 million
Media	2 TB	.5 TB
Languages	33	18
Publications	3500+	300
Users	3200	600

The Core Idea

- Human communication is a single unified process.
- The integration occurs through competition at the moment of speech.
- The time scales of the processes vary across 7 major spatio-temporal frames, each with varying components.

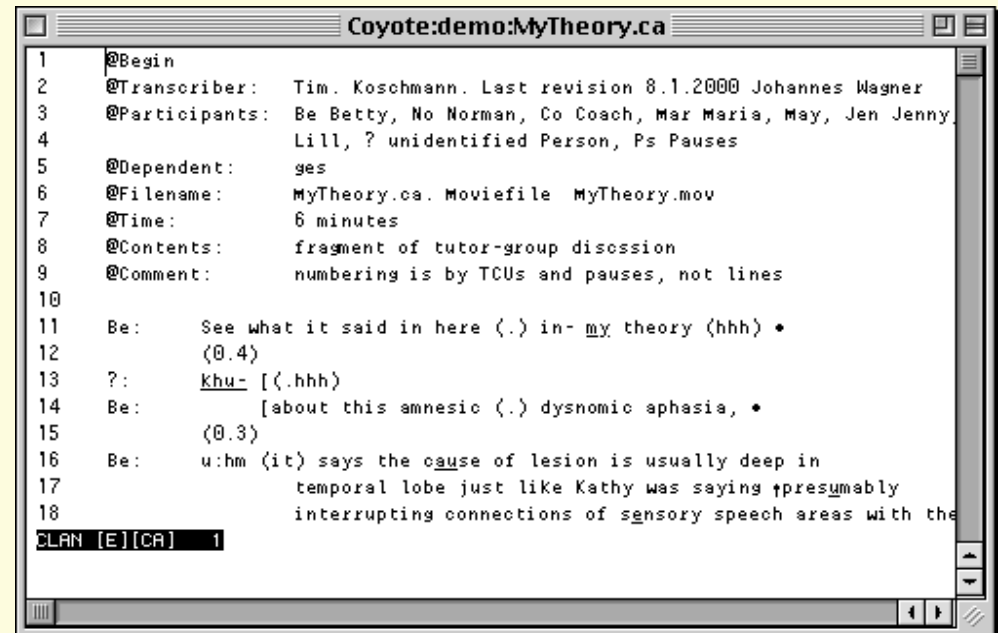
7 spatio-temporal frames

1. Production
2. Perception
3. Interaction
4. Social affiliation
5. Development
6. Diachrony
7. Phylogeny

Integration and Capture

- All of the space-time frames must show their effects and be conditioned in the current moment in time and space.
- We can capture the current moment and current place on video.
- However, we will need to compare across time and space to understand the textures of the competing component processes.

A sample moment: Transcript linked to video



Other views

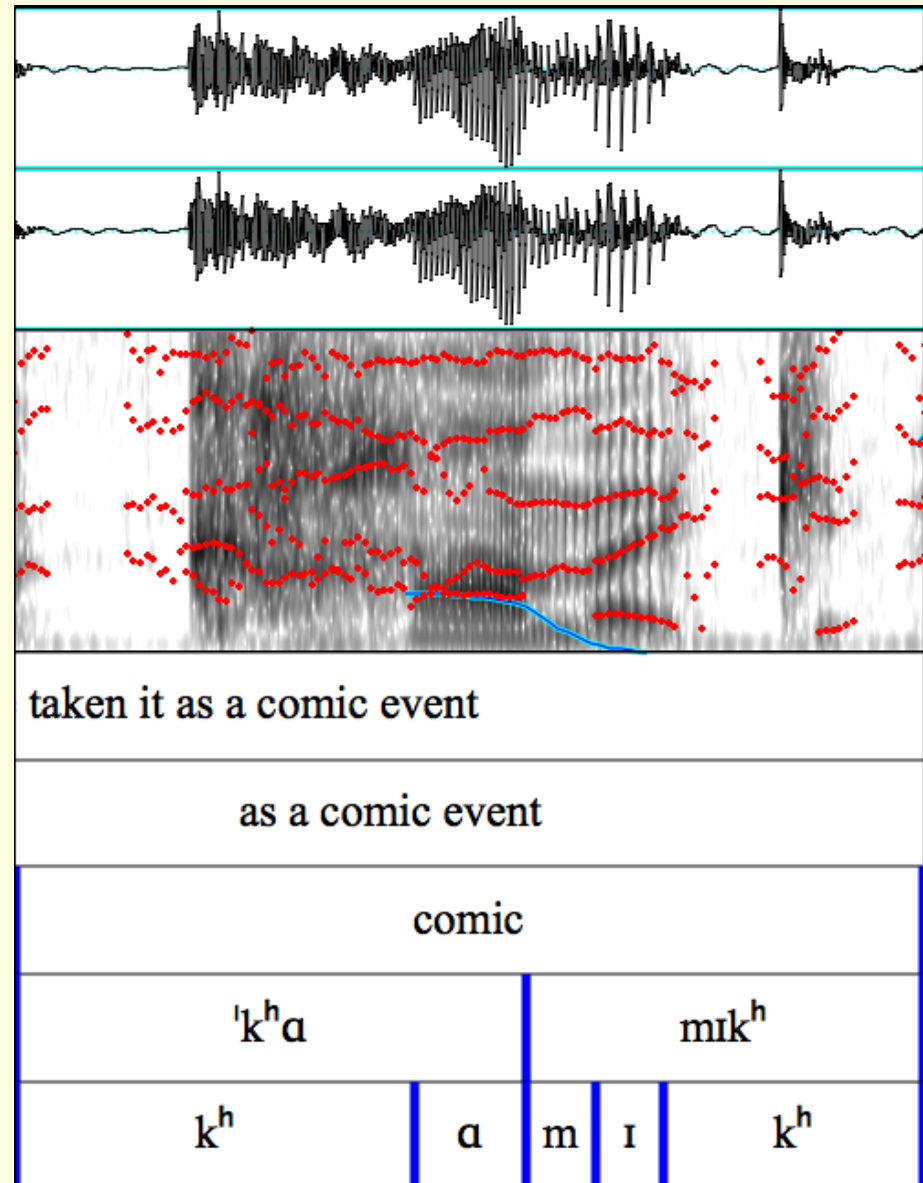
The screenshot displays the Elan software interface for the file 'mytheory.eaf'. The top menu bar includes File, Edit, Annotation, Tier, Type, Search, View, Options, Window, and Help. The main window is divided into several sections:

- Video View:** A small window on the left shows a video of a person sitting at a desk with a large chart of brain scans in the background.
- Controls:** On the right, there are sliders for Volume (0 to 100) and Rate (0 to 200), both currently set to 100. Above these are tabs for Grid, Text, Subtitles, and Controls.
- Timeline and Playback:** Below the video, a timeline shows the current time at 00:00:40.530. A selection range is marked from 00:00:40.530 to 00:00:41.515. Playback controls (stop, previous, play, next, full screen, etc.) are visible.
- Annotation Tiers:** The bottom section shows a list of annotation tiers on the left and their corresponding time-aligned text on the right:
 - *BET
 - *UNK
 - *NOR
 - *COA
 - %gpx@NOR
 - *MAR: if you lift up +/., that little temporal lobe | insid | #0 | Middle top? | 0.
 - %gpx@MAR: brings R hand in | lifts R hand above head | Maria poin
 - %gpx@COA: Points with R hand from seat t

Deeper Views



Still Deeper Views



Gestural Views

Segment	N1
Action	rests chin on hand, elbow on table, right shoulder back
Gaze	front to Deedee
Classification	Attention
Meaning	Attention

*D: 「så er det snart」 「torturtid→」

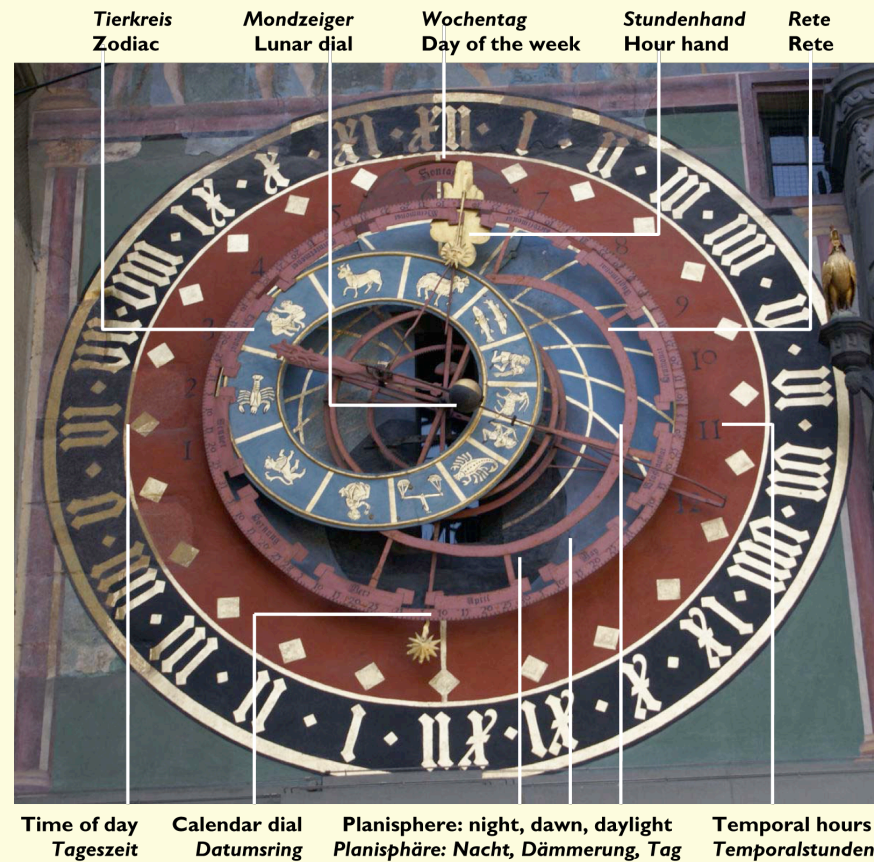
%ges: 「-----D1-----」 「----D2----」
 「-----N1-----」

%com: assimilating the pronunciation of a danish actor in a
 then tv show

pic *

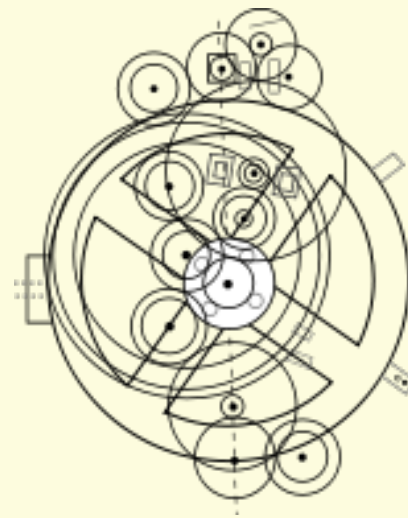
pic *

Meshing of space-time scales



Orloj of Prague -- 1490

The Antikythera – Greece 150BC



1. Production Wheels

- lexical access (PDP)
- phonological activation (Dell)
- morphological combination (MacWhinney)
- gang effects (all six linguistic levels)
- rote, combination (puzzle, puddle)

2. Perceptual Wheels

- statistical learning (Saffran)
- attention to ends vs beginnings (Juszczyk)
- attention to stress (Mehler)
- uptake vs input (SLA literature)
- input vs output frequency
- changes in attentional biases based on input
- analytic vs gestalt packaging (Peters)

3. Interactional Wheels

- Child self correction, retrace (Dialectic Model)
 - Parental correction, recast (Bohannon, ...)
 - Variation sets, scaffolding (Ochs, Waterfall)
 - Repetition, imitation, choral (Forrester)
 - Topic maintenance (TBA)
 - Turn projection, completion, overlap (TBA)
- (all of these with phonological and gestural signals!)

4. Social-affiliative wheels

- Gaze contact (Tomasello)
- Affiliation particles (Bill Wells)
- Body alignment (Zlatev)
- Disaffiliation, breakdown
- Fine-tuning (Snow, Sokolov)
- Perspective taking – narrative, discourse (MacWhinney, Morgenstern)

5. Developmental Wheels

- OT constraints
- Growth of motor control – Oller, Davis
- Entrainment, coupling – Thelen
- Physical changes in vocal tract
- Entrenchment, neural commitment (Kuhl)

6. Diachronic Wheels

- Uniformism – Grimm's Law
- Northern Cities shift, push-pull
- Lexical diffusion (Ota)
- Founder's effect
- Long-term social-affiliation (Labov)

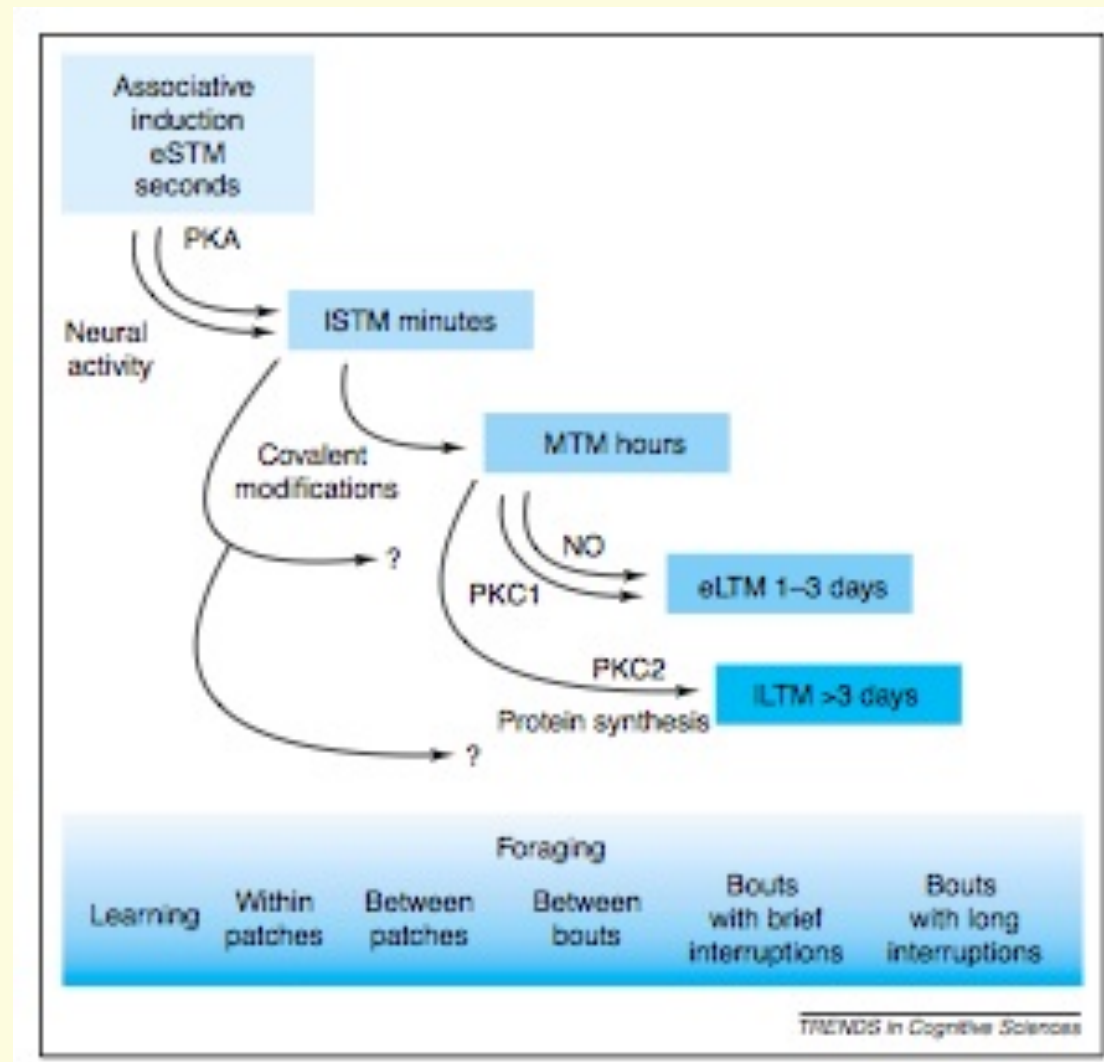
7. Phylogenetic Wheels

- CV frame-content (Davis-Macneilage)
- Articulatory control (FoxP2)
- SLI connectivity (Evans)

Memory Reflexes of Frames

- short-term precise acoustic
- featural longer storage
- short-term lexical
- growing lexical
- hippocampal reentrant consolidation
- proceduralization
-

Timeframes in memory in bees



Linking Timeframes

- Frames impact memory which then provides inputs to the competition
- Slower, marked processes must come to override initial, unmarked processes
- Competition Model: Effects of frequency, reliability, availability, detectability, conflict validity (Stemberger), error tagging

Examples

- Getting rid of final fortis
- Getting the right Zapotec diminutive
- Puzzle-puddle competition
- Frequency when two paths possible (Levelt)
- Merging lexicon with segmentation algorithms (JCL special issue)
- Islands, leading forms

The key tests

- Modeling details of individual differences
- Linking the differences to measurable other factors
- Modeling markedness
- Eventually, linking to L2 acquisition

Rich Data Needed

- Many children
- Dense corpora
- Detailed transcriptions
- Planned variations (experiments)

Data-sharing

- Clearly defined formats can be equated
- Interoperability of program formats
- 42 reasons not to share data
- The reason to share
- The solutions:
 - Methods for password protection
 - Methods for anonymization
 - Credit to contributor
 - Group commitment

Analysis Methods

1. Bag of Words
2. QDA = a.k.a. Hand Coding
3. Tagging = a.k.a. Automatic Coding
4. Profiles = a.k.a. Canned Analyses
5. Group/treatment comparisons
6. CA Analysis
7. Gesture Analysis
8. Phonetic Analysis
9. Collaborative Commentary
10. Error analysis
11. Longitudinal analysis
12. Modeling

1. Bag of Words (BoW)

- Basic method of Corpus Linguistics
- For written data, there are many many resources: Google, BNC, Libraries, LDC
- But for spoken data, TalkBank is the major open source
- Core BoW analyses support
 - Usage-based learning models in L1 and L2
 - Theories in eight other areas

BoW Methods

- Basic Programs (CLAN and BNC)
 - **FREQ** (BNC links to t-tests) / **STATFREQ**
 - **KWAL** with windows
 - **COMBO** (regular expressions)
- **WebCLAN** (limited)
- Download and run locally
- **X-Query Search Engine** (in preparation)

BoW Methods

- **FREQ -> STATFREQ -> EXCEL**

	over	ow	own	paper	papers	people	pig	pillow
o	0	0	0	1	0	8	0	0
.	0	1	0	1	0	1	1	0
y	0	0	1	0	1	1	0	0
y	2	2	0	1	0	0	0	1

- **KWAL -> clickable output**
- **Limiting through GEM**
 - @Bg: conversation ending
 -
 - @Eg: conversation ending

2. Qualitative Data Analysis (QDA) = Coding

1. Build Coding System
2. Use Coder's Editor to insert codes
3. Use RELY to compare coder accuracy
4. RELY output pinpoints disagreements
5. Click and play disagreements to refine coding system

Examples: Rollins INCA, MUMIN in Anvil

Speech Act Coding

```
@Begin
@Languages: en
@Participants: MOT Mother, CHI David Target_Child
@ID: en|rollins|MOT||||Mother||
@ID: en|rollins|CHI|1:8.||||Target_Child||
@Activities: book

*MOT: ahhah: look we can read books Tim .
%spa: $DHA:YY $DHA:RP

*MOT: it's a look and see <book> [>] .
%spa: $DHA:ST

*MOT: <ahhah> [>] we open it up and there are a set of eyes
%spa: $DJF:ST $DHA:ST

*MOT: <the bear has a baby> [>] bottle .
%spa: $DHA:ST

*MOT: yes # David has baby <bottles> [>] .
%spa: $DRP:ST

*MOT: <oh> [>] .
%spa: $DHA:MK

*MOT: <there's a mirror> [>] .
%spa: $DJF:ST

*MOT: can David see <David> [>] .
%spa: $DHA:RQ

*CHI: 0 .
%act: CHI look-s at rattle in hand then puts rattle in mouth
```

http://xml.talkbank.org:8888/talkbank/rollins/

[0:13:19 - 0:19:19]

*MOT: <there's a mirror>

%spa: \$DJF:ST

Orange Blue

User Comment 1: A good coding sequence with high correlation to the research community

User Comment 2: I refer the reader to Keck et. al where you will see that this coding scheme is highly disputed.



QDA through Naked Video

- Terabytes of video
 - Speechome, Classroom, Resident Care
- No transcripts
- Occasional sign posts
- Sparse speech recognition
- Automatic video analysis

3. Tagging

- Morphosyntax – MOR, POST
 - 12 languages
 - Some languages need more training
 - With correct transcription, accuracy is at 98%
 - MOR generates tags
 - POST disambiguates
 - POSTMORTEM examines residual issues

Tagging (cont.)

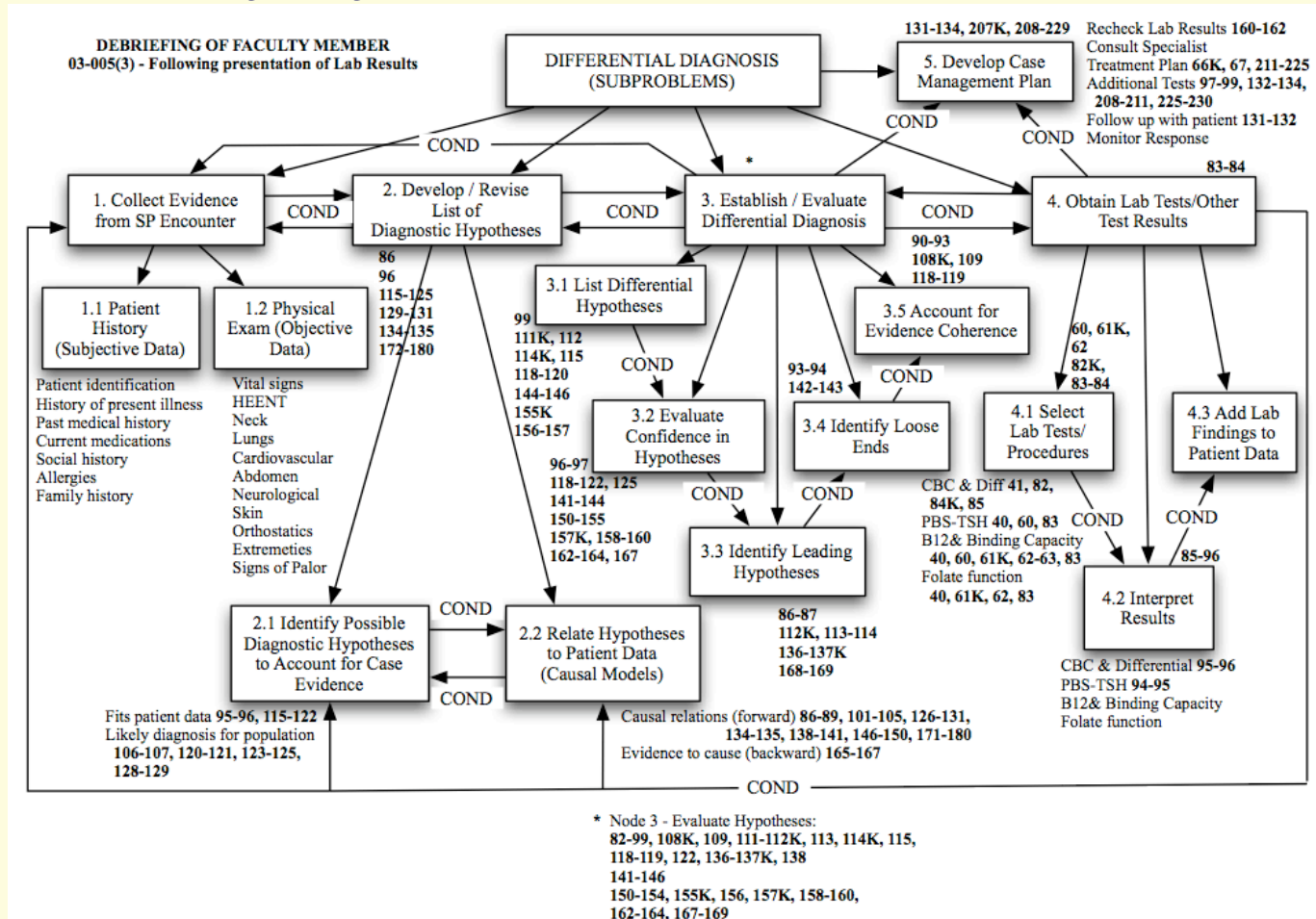
- GRASP uses output of MOR to add grammatical relation (GR) dependency structure with 38 relations.
 - English, Japanese, Hebrew, Spanish
 - Accuracy is at 93%, more work still needed
- Tagging for CA categories?
 - Eckhardt, Mondada, & Wagner

Searchable Features

Cutoffs	+/.
Overlaps	┌ ┐ └ ┑
Fillers	um, em
Pauses, pause length	(.) (6.2) or #6_2
Repeats, retraces	[/] [//]
Prosodic	↑ ↓ ↘ ↗
Latching	≈ +,
Paralinguistic	&=
Others	— —

Propositional Tagging

Polycythemia - Frederiksen

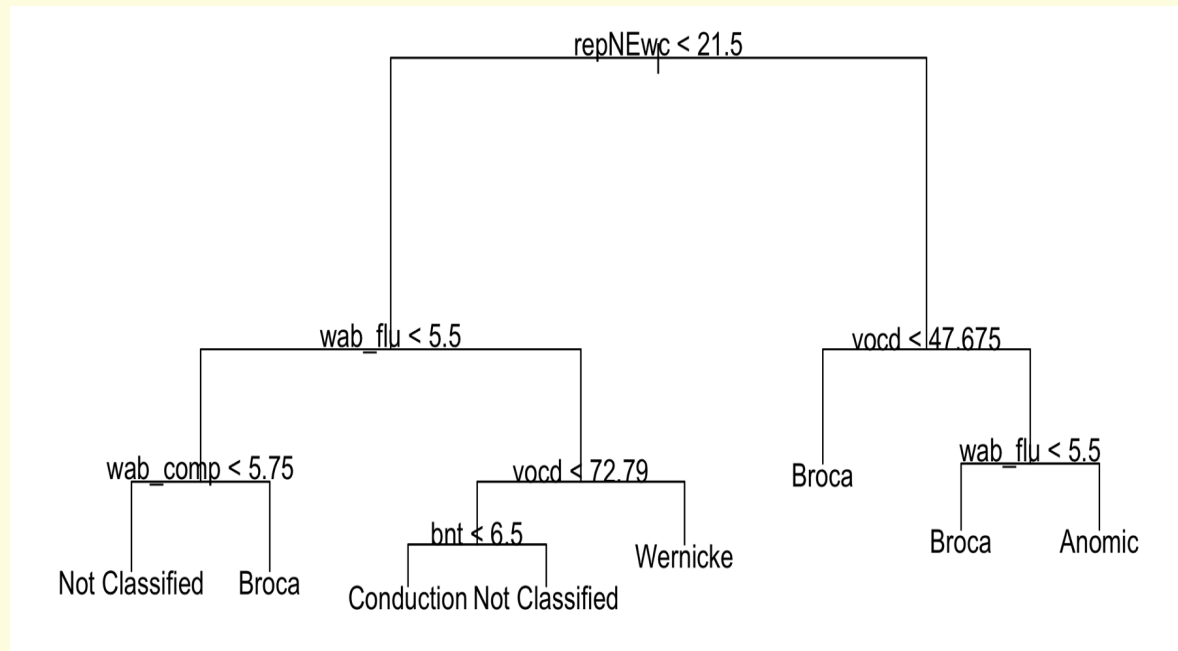


4. Language Profiles

- Phonological inventories, TAKI
- DSS (English, Japanese)
- IPSyn
- MORTABLE
 - Parts of speech

◇	A	B	C	D	E	
1	speaker ID	*:wh	adj,adj:*	adv,adv:*	aux,aux:*	cn
2	eng Elman PAR Conduction elman01a Participant	12	27	92	50	
3	eng Elman INV elman01a Investigator	10	14	27	17	
4	eng Elman PAR Conduction elman02a Participant	14	37	55	33	
5	eng Elman INV elman02a Investigator	5	16	27	16	
6	eng Elman PAR Broca elman03a Participant	12	14	30	17	
7	eng Elman INV elman03a Investigator	7	14	21	10	
8	eng Elman PAR NotAphasicByWAB elman04a Pa	13	28	69	44	
9	eng Elman INV elman04a Investigator	8	31	29	18	
10	eng Elman PAR Anomic elman05a Participant	6	15	43	34	

AphasiaBank Classification



Clinician Types by K-means clusters

	1	2	3	4	5	6
Anomic	3	11	0	1	5	0
Broca	4	2	1	8	3	5
Conduction	4	0	0	1	0	4
Global	0	0	3	0	0	0
Not Classified	0	1	2	2	3	1
Other	0	0	0	1	0	0
TCM	1	0	0	0	0	0
Wernicke	1	0	0	1	0	1

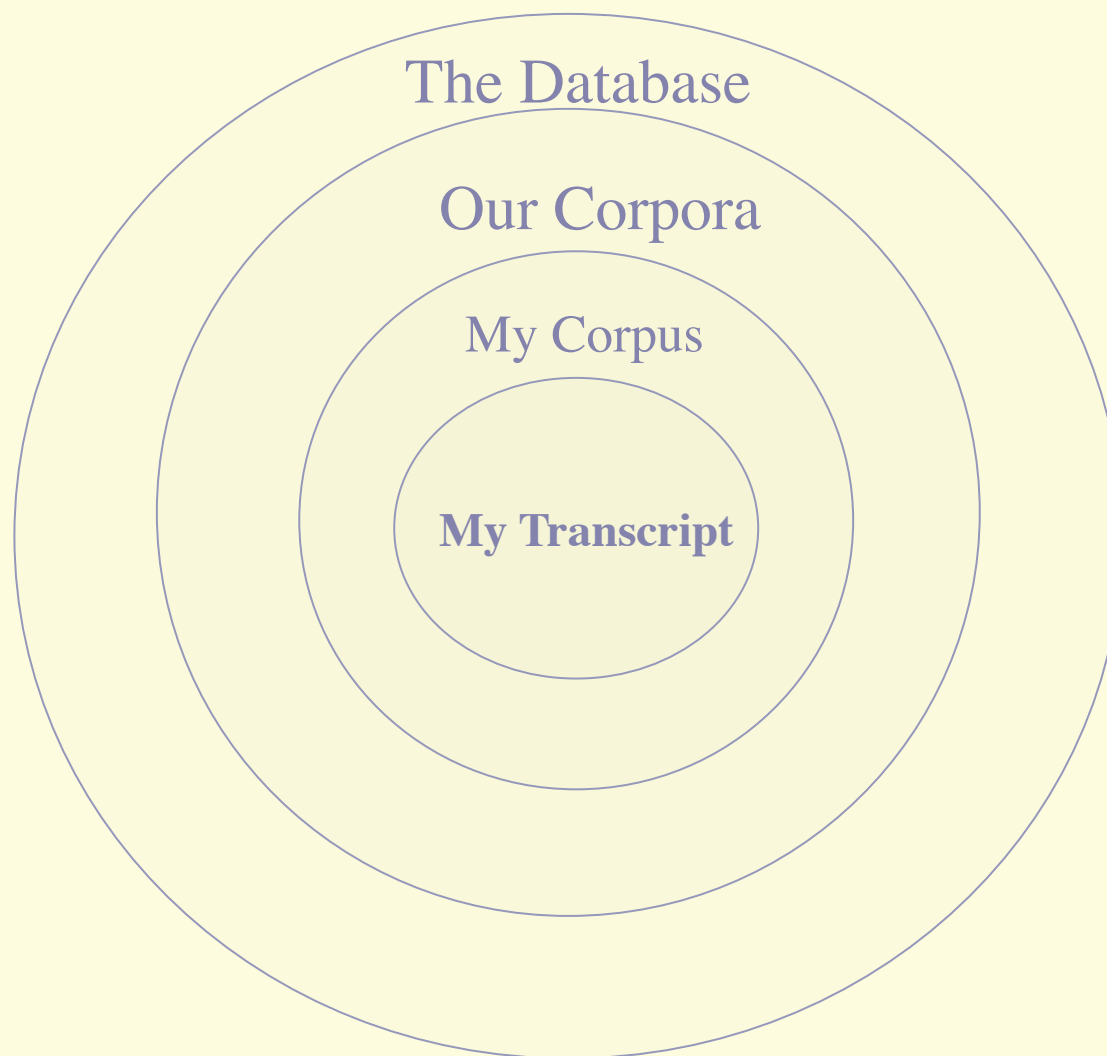
5. Group Comparisons

- Pretest – Treatment – Posttest
 - Measure gain scores – AphasiaBank Wright
 - L2 increases in fluency (Praat and TIMEDUR) from 4/3/2 training – Nel de Jong
 - Classroom discourse
 - Accountable discourse
 - MacWhinney and Arkenberg
 - Lauren Resnick, Beth Warren, Sarah Michaels

6. CA Analysis

- CA Database
 - SamtaleBank (CALPIU?)
 - STEM/L2 classroom data
 - Newport Beach, Watergate, CallFriend
 - Koschmann Competency
 - Santa Barbara

CA Corpora?



CA Tools

- Overlap alignment through CAFont and INDENT
- Removal of constraints on sentences, focus on TCUs and turns
- Line numbers on and off
- Alignment to audio – sonic CHAT
- Special characters


<u>Character Name</u>	<u>Char</u>	<u>Function</u>
up-arrow	↑	shift to high pitch
down-arrow	↓	shift to low pitch
double arrow tilted up	↗	rising to high
single arrow tilted up	↗	rising to mid
level arrow	→	level
single arrow tilted down	↘	falling to mid
double arrow down	↘	falling to low
infinity mark	∞	unmarked ending
double wavy equals	≈	latching≈ or +≈latch
triple wavy equals	≈	+≈ text
triple equal	≡	≡uptake
raised period	·	inhalation
open bracket top	[top begin overlap
close bracket top]	top end overlap
open bracket bottom	⌊	bottom begin overlap
closed bracket bottom	⌋	bottom end overlap
up triangle	Δ	ΔfasterΔ
down triangle	∇	∇slower∇
low asterisk	*	*creaky*

7. Gesture Analysis

- Detailed tiers in ANVIL – MUMIN, FORM
- Basic time linkage in Elan – HKSL
- Automatic interoperability between ANVIL, Elan, and CLAN
- Microscopic zooming in CLAN
 - Links to “sequence” subfiles
 - Links to “snapshot” subfiles

In CHAT and CLAN

Movie - Sound



1425

/ \

0 3200

Repeat b0 msec

MYTHEORY.MOV

5/22/04

Coyote:demo:MyTheory.ca

```

1  @Begin
2  @Transcriber:  Tim. Koschmann. Last revision 8.1.2000 Johannes Wagner
3  @Participants: Be Betty, No Norman, Co Coach, Mar Maria, May, Jen Jenny,
4                  Lill, ? unidentified Person, Ps Pauses
5  @Dependent:    ges
6  @Filename:     MyTheory.ca. Moviefile  MyTheory.mov
7  @Time:         6 minutes
8  @Contents:     fragment of tutor-group disocssion
9  @Comment:      numbering is by TCUs and pauses, not lines
10
11  Be:    See what it said in here (.) in- my theory (hhh) •
12         (0.4)
13  ? :    khu- [(.)hhh)
14  Be:    [about this amnesic (.) dysnomic aphasia, •
15         (0.3)
16  Be:    u:hm (it) says the cause of lesion is usually deep in
17          temporal lobe just like Kathy was saying ↑presumably
18          interrupting connections of sensory speech areas with the
CLAN [E][CA] 1
    
```

In ELAN

The screenshot displays the ELAN software interface for the file 'mytheory.eaf'. The top menu bar includes File, Edit, Annotation, Tier, Type, Search, View, Options, Window, and Help. Below the menu is a toolbar with icons for navigation and editing. The main window is divided into two panes. The left pane shows a video of a person sitting at a desk, pointing at a board with several brain diagrams. The right pane contains controls for Volume (0 to 100) and Rate (0 to 200). Below these controls is a timeline with a selection bar indicating the current time range (00:00:40.530 to 00:00:41.515). The bottom pane shows a list of annotations with their corresponding time ranges and text. The annotations are organized into tiers, with some tiers highlighted in color (e.g., *BET in red, *UNK in blue, *NOR in green, *COA in yellow). The annotations include text such as 'if you lift up +/- that little temporal lobe', 'brings R hand in lifts R hand above head', and 'Points with R hand from seat t'.

Elan - mytheory.eaf

File Edit Annotation Tier Type Search View Options Window Help

Grid Text Subtitles Controls

Volume: 100

Rate: 100

00:00:40.530 Selection: 00:00:40.530 - 00:00:41.515 985

Annotations:

- *BET
- *UNK
- *NOR
- *COA
- %gpx@NOR
- *MAR: if you lift up +/- that little temporal lobe | insid | #0 | Middle top? | 0.
- %gpx@MAR: brings R hand in lifts R hand above head | Maria poin
- %gpx@COA: Points with R hand from seat t

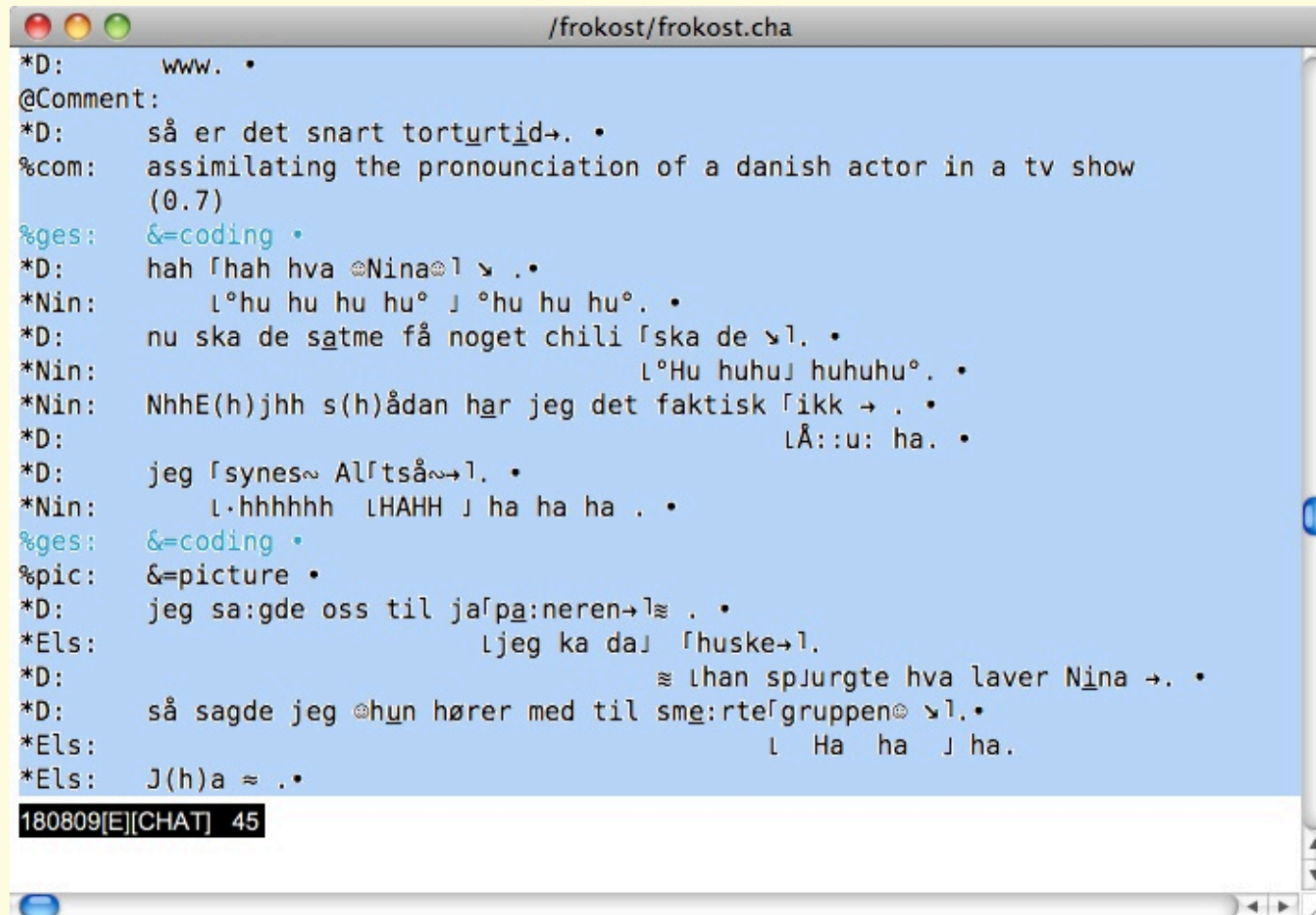
Torturtid



5/22/04

50

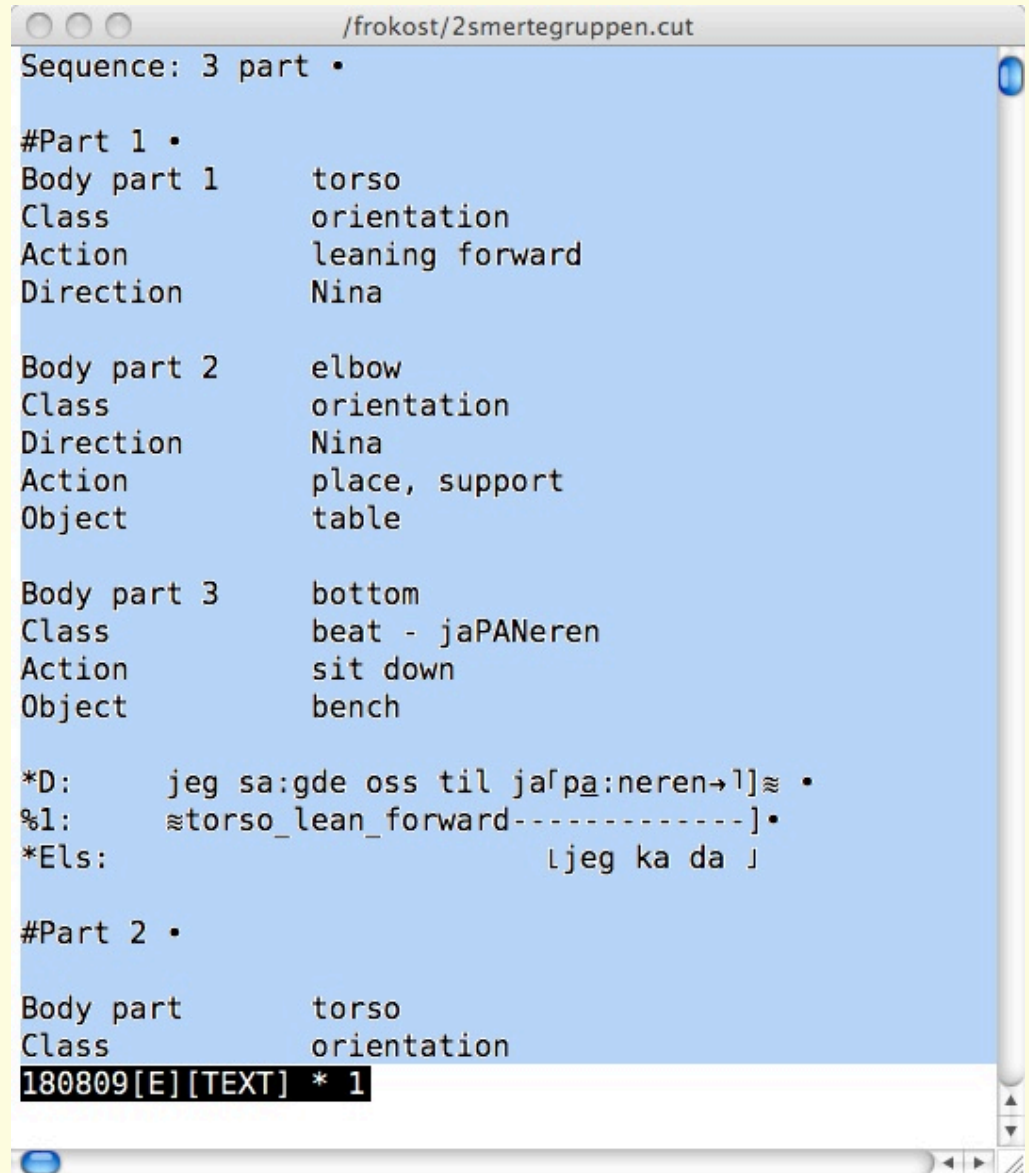
Overall transcript



```
*D:      www. •
@Comment:
*D:      så er det snart torturtid→. •
%com:    assimilating the pronunciation of a danish actor in a tv show
(0.7)
%ges:    &=coding •
*D:      hah [hah hva @Nina@] v. •
*Nin:    l°hu hu hu hu° ] °hu hu hu°. •
*D:      nu ska de satme få noget chili [ska de v]. •
*Nin:    l°Hu huhu] huhuhu°. •
*Nin:    NhhE(h)jhh s(h)ådan har jeg det faktisk [ikk →. •
*D:      LÅ::u: ha. •
*D:      jeg [synes~ Al[tså~]. •
*Nin:    l.hhhhhh [HAHH ] ha ha ha . •
%ges:    &=coding •
%pic:    &=picture •
*D:      jeg sa:gde oss til ja[på:neren→]≈ . •
*Els:    [jeg ka da] [huske→].
*D:      ≈ lhan spjurgte hva laver Nina →. •
*D:      så sagde jeg @hun hører med til sme:rte[gruppen@ v]. •
*Els:    l Ha ha ] ha.
*Els:    J(h)a ≈ . •
180809[E][CHAT] 45
```

Sequence Subfiles

- Three parts
- Each part has components
- Each part linked
- Each part displayed



```
/frokost/2smertegruppen.cut
Sequence: 3 part •

#Part 1 •
Body part 1      torso
Class            orientation
Action          leaning forward
Direction        Nina

Body part 2      elbow
Class            orientation
Direction        Nina
Action          place, support
Object          table

Body part 3      bottom
Class            beat - jaPANeren
Action          sit down
Object          bench

*D:      jeg sa:gte oss til ja[pa:neren→l]≈ •
%1:      ≈torso_lean_forward-----]•
*Els:                                Ljeg ka da J

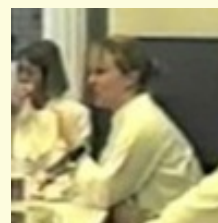
#Part 2 •

Body part      torso
Class          orientation
180809[E][TEXT] * 1
```

Snapshot Files

*C: (0.2) participants ↗ ↓ * ɾo:ɾɿ *

%vis: -----1-----|-----2-----|-----3--



1. on uttering the syllable "ci", C reaches for a pencil with her right hand and paper with left hand.
2. On uttering the syllable "pants", C grabs a pencil with right hand and the paper with left hand.
3. On "or", she lifts the paper from the table.

8. PHON \leftrightarrow CLAN \leftrightarrow Praat

Orthography

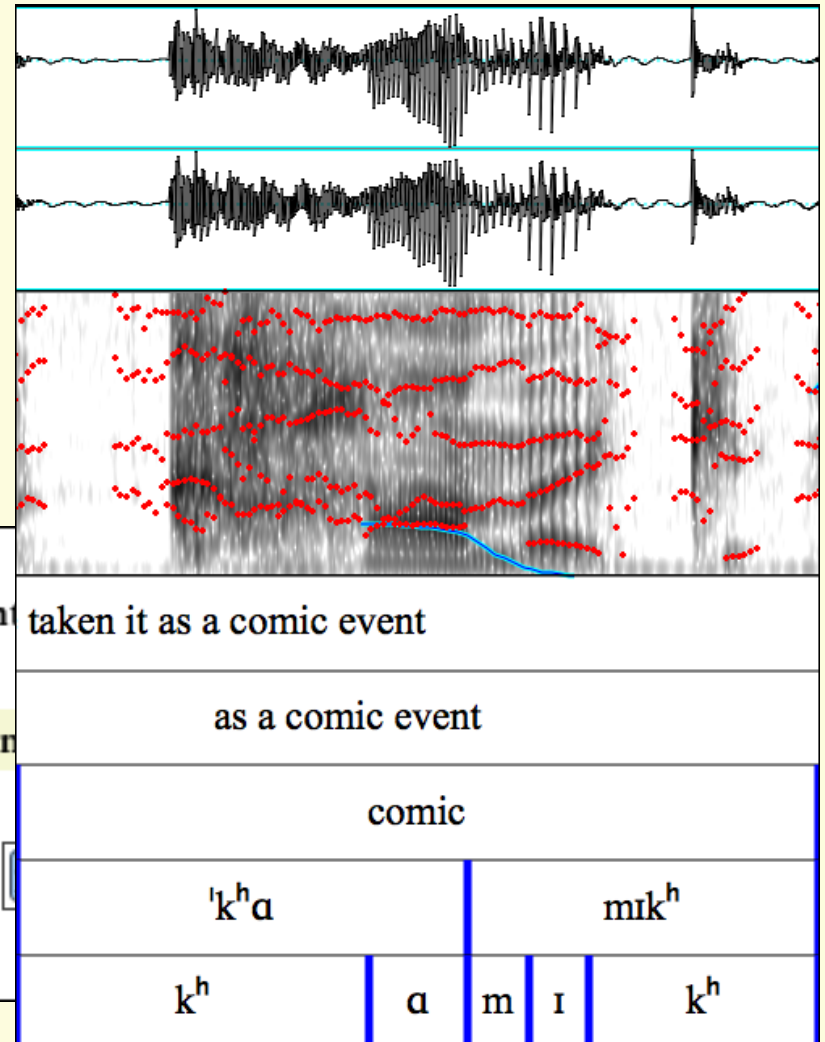
[You know] [a lot of people] [have taken it] [as a comic event]

IPA Actual

[jə 'nou] [ə 'lɑr əv 'pʰi:pəl] [əv 'tʰeɪkən 'ɪ?] ['æz ə 'kʰɑmɪkʰ ɪ'ven]

Actual Syllables

j	ə	n	o	u	ə	l	ɑ	r	ə	v	pʰ	i:	p	ə	l	ə	v	tʰ	e	ɪ	kən	ɪ	?	æ	z	ə	kʰ	ɑ	m	ɪ	kʰ	ɪ	v	ɛ	n	tʰ
---	---	---	---	---	---	---	---	---	---	---	----	----	---	---	---	---	---	----	---	---	-----	---	---	---	---	---	----	---	---	---	----	---	---	---	---	----



Phonetic Data

Acquire Vowel Data

Please select the acoustic parameters to import from Praat

☒ Duration
☒ Pitch
 [...etc...]

☐ Intensity
 [...etc...]

☒ Formant Structure
☒ F1 ☒ F2 ☐ F3 ☐ F4

Sample Location: ☒ Vowel Midpoint
☐ % from start
☐ % from end

[...etc...]

Set Praat Settings

Set as Defaults

Acquire Data

Session Name:	Corpus for Praat.Praat session	
Session Date:	2008-05-27	
Participants:	Name	Age
	Arthur	3;6.4
Report Type:	Vowel Data	
Utterance:	[You know] [a lot of people] [have taken it] [as a comic event]	
Word Group:	[as a comic event]	
Word:	comic	
Word IPA:	'kʰamɪkʰ	
Vowel:	a	
Duration (ms):	105	
Intensity (dB):	MidPoint:	74
	25% from Start:	71
	25% from End:	75
Pitch (Hz):	MidPoint:	185
	25% from Start:	188
	25% from End:	177
F1:	MidPoint:	902
	25% from Start:	919
	25% from End:	862
F2:	MidPoint:	1308
	25% from Start:	990
	25% from End:	1239

9. Collaborative Commentary

23 *MOT: you need a little help there I think, be careful . ▶
24 *MOT: miss athlete .
25 *CHI: athlete . ▶
Add Comment
macw posted this comment on May 26th, 2010 5:31 pm
\$PHO remarkably accurate pronunciation
26 %xpho: 'æpəlɪt
27 *MOT: 'æθəlɪt@u [: athlete_a] .
28 *MOT: had a very strong accent, what about (macw@: [macw@: 1.2

Comment Tagging, Filtering

- Automatic: author, date, media begin-end
- Author self-characterized metadata (role, faction, position, credentials)
- Commentary type (refutation, defense, elaboration, analogy, statistics, case law, gesture-speech match)
- Filters: only teacher, only from colleagues, etc.

10. Error Analysis

- Basic to work in CHILDES, BilingBank, and AphasiaBank
- Main line coding system
 - goed [: went] [* +ed]
 - I want 0to go home.
- Complete system for aphasia, speech errors

11. Sequential Analysis

- Variation sets, recasting, CHIP, fine-tuning
- If CDS has “want X”, does child increase use of “want go home”
- Code sequences through CHAINS and KeyMap
- Phonological Model-Replica analysis
- Richer analysis through MacShapa

12. Modeling

- Neural networks
 - Some (PDP, MOSAIC) just use rough counts
 - DISLEX uses actual CDS from CHILDES
 - Segmentation models use Brent corpus as the gold standard for input
 - Most recent models take the auditory form of the CDS as input for learning
 - Eventually, models will induce from complete multimedia databases (Speechome)

Conclusions

- We can transform the study of conversation
- But we still need to provide the technical basis for data-sharing, interoperability, and collaborative commentary
- After that, the major barrier is a full commitment to data-sharing
- And patience to integrate across seven time scales.

The Databases of the Future

- Individuals as keys
- Institutions as keys
- Activities as keys
- Linking: psychology, linguistics, sociology, political science, economics, genetics, genomics, biology, geography, and anthropology'