

# Talk 12: Speech, Analysis and Linguistics

James Cummings

July 2009

# Spoken Texts

A spoken text may contain any of the following components:

- utterances
- pauses
- vocalized but non-lexical phenomena such as coughs
- kinesic (non-verbal, non-lexical) phenomena such as gestures
- entirely non-linguistic incidents occurring during and possibly influencing the course of speech
- writing, regarded as a special class of incident in that it can be transcribed, for example captions or overheads displayed during a lecture
- shifts or changes in vocal quality

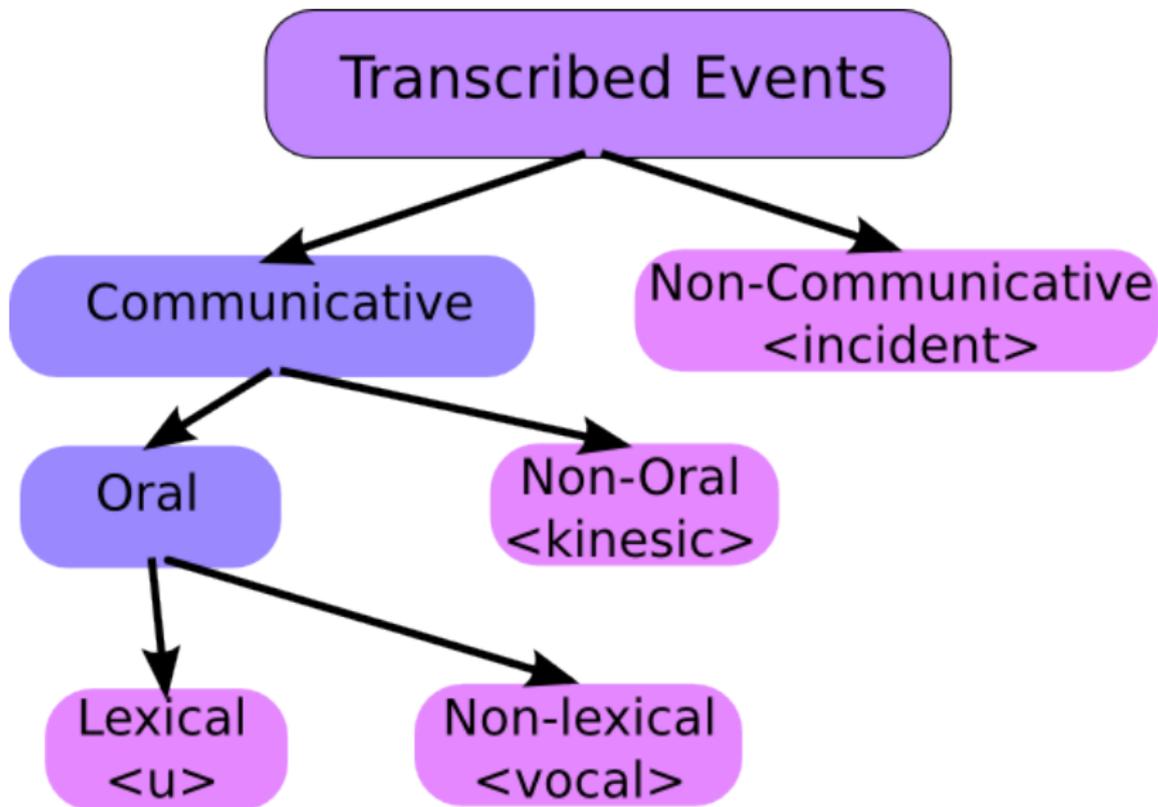
## TEI for spoken texts

The spoken texts module proposes an XML-based markup for

- a lexically useful subset of speech phenomena
- a rich set of associated contextual information (metadata)
- linking and alignment mechanisms

As a part of the TEI scheme, it can be integrated, extended, and customized in a standard way.

## What is "lexically useful"?



## The notion of "utterance"

- problematic, but pragmatic
- a sequence of speech from a single speaker
- may be grouped into higher-level `<div>`s
- or fragmented into smaller segments `<seg>` or `<s>`
- the `@who` attribute points to speaker information

# Transcribed Speech

**Elements defined:** <broadcast>, <equipment>, <incident>, <kinesic>, <pause>, <recording>, <recordingStmt>, <scriptStmt>, <shift>, <u>, <vocal>, <writing>,

**Classes defined:** att.duration, model.divPart.spoken, model.global.spoken, model.recordingPart

## Simple examples

Mixture of utterance and 'paralinguistic' information:

```

<u who="#Jan">This is just delicious</u>
<incident>
  <desc>telephone rings</desc>
</incident>
<u who="#Kim">I'll get it</u>
<u who="#Tom">I used to <vocal>
  <desc>coughs</desc>
</vocal> smoke a lot</u>
<u who="#Bob">
  <vocal>
    <desc>sniffs</desc>
  </vocal>He thinks he's tough
</u>
<vocal who="#Ann">
  <desc>snorts</desc>
</vocal>
<u who="#Tom">Yeah
<kinesic>
  <desc>gives uplifted middle finger sign</desc>
</kinesic>
</u>

```

# Back channelling

```
<u who="#a">So what could I have done <vocal who="#b">  
  <desc>tut-tutting</desc>  
</vocal> about it anyway?</u>
```

## Example using other TEI elements

```

<u who="#mar">you never <pause/> take this cat for
show and tell
<pause/> meow meow</u>
<u who="#ros">yeah well I dont want to</u>
<incident>
  <desc>toy cat has bell in tail which continues
  to make a tinkling sound</desc>
</incident>
<u who="#ros">because it is so old</u>
<u who="#mar">how <choice>
  <orig>bout</orig>
  <reg>about</reg>
</choice>
<emph>your</emph> cat <pause/>yours is <emph>new</emph>
<kinesic>
  <desc>shows Father the cat</desc>
</kinesic>
</u>
<u trans="pause" who="#fat">thats <pause/> darling</u>
<u who="#mar">no <emph>mine</emph> isnt old
mine is just um a little dirty</u>

```

## Shifts in voice quality

- Classic multiple hierarchy problem
  - can use `<shift>` or `<milestone>` to mark boundaries...
  - ... or can use typed `<seg>` elements
- useful also for code shifting

```
<u who="#LB">  
  <shift feature="loud" new="f"/>Elizabeth  
</u>  
<u who="#EB">Yes</u>  
<u who="#LB">  
  <shift feature="loud"/>Come and try this <pause/>  
  <shift feature="loud" new="ff"/>come on  
<shift feature="loud"/>  
  <milestone type="kreol" unit="codeshift"/>tinva  
</u>
```

## Sample prosodic feature list

(based on Boase, Survey of English Usage, 1990)

---

tempo	(fast, slow, getting faster, slower, etc.)
loud	loud, soft, getting louder, slower
pitch range	high, low, wide, narrow, ascending...
tension	slurred, tense, staccato, legato...
rhythm	regular, irregular, spiky rising or falling...
voice quality	whisper, husky, falsetto, giggle, sobbing, yawning, sighing...

Researchers need to define their own terms

## <shift/> example

`<u who="#a">`Listen to this `<shift new="reading"/>`The government is confident, he said, that the current economic problems will be completely overcome by June`<shift/>` what nonsense`</u>`

or as an <incident>

```
<u who="#a">Listen to this  
<incident>  
  <desc>reads aloud from newspaper</desc>  
</incident> what nonsense</u>
```

# <vocal> vs <u>

Compare:

```
<vocal who="#ann">  
  <desc>snorts</desc>  
</vocal>
```

and

```
<u who="#ann">  
  <vocal>  
    <desc>snorts</desc>  
  </vocal>  
</u>
```

## <writing> example

```
<u who="#a">look at this</u>  
<writing who="#a" type="newspaper" gradual="false">  
Government claims economic problems <soCalled>over by June</soCalled>  
</writing>  
<u who="#a">what nonsense!</u>
```

## Timing issues

- pausing: use `<pause>` element
- duration: use `@dur` attribute
- synchronization: use `@synch` attribute
- overlap: use `@trans` attribute

## <pause e> example

<u>Okay <pause dur="PT2M"/>U-m<pause dur="PT75S"/>the scene opens up  
<pause dur="PT50S"/> with <pause dur="PT20S"/> um <pause dur="PT145S"/> you  
see  
a tree okay?</u>

# Overlap

```
Mutt: Have you heard the --  
Jeff: the election result?  
Mutt: It's a disaster!
```

```
<u who="#mutt">have you heard the</u>  
<u trans="latching" who="#jeff">the election result</u>  
<u who="#mutt">its a disaster</u>  
<u who="#jeff" trans="overlap">its a miracle</u>
```

## More overlap

```
<u who="#tom">I used to smoke <anchor xml:id="TS-p10"/> a lot more than this  
<anchor xml:id="TS-p20"/> but I never inhaled the smoke</u>  
<u start="#TS-p10" end="#TS-p20" who="#bob">You used to smoke</u>
```

# Synchronization

```
<u who="#mutt">have you heard <anchor synch="#t1"/>the</u>  
<u who="#jeff" synch="#t1">the election result</u>  
<u who="#mutt" synch="#t2">its a disaster</u>  
<u who="#jeff" synch="#t2">its a miracle</u>  
<!-- Elsewhere in Document -->  
<timeline origin="#t1">  
  <when xml:id="t1"/>  
  <when xml:id="t2"/>  
</timeline>
```

## <timeline> example

```
<timeline unit="s" origin="#TS-P1">
  <when xml:id="TS-P1" absolute="12:20:01"/>
  <when xml:id="TS-P2" interval="4.5" since="#TS-P1"/>
  <when xml:id="TS-P6"/>
  <when xml:id="TS-P3" interval="1.5" since="#TS-P6"/>
</timeline>
```

```
<u xml:id="TS-U1" start="#TS-P2" end="#TS-P3">This is my
<anchor synch="#TS-P6" xml:id="TS-P6A"/> turn</u>
```

The start of utterance TS-U1 is aligned with TS-P2 and its end with TS-P3. The transition between the words my and turn occurs at point TS-P6A, which is synchronous with point TS-P6 on the timeline.

# Using elements seen elsewhere

```
<u>  
  <del type="truncation">s</del>see  
  <del type="repetition">you you</del> you know  
  <del type="falseStart">it's</del> he's crazy  
</u>
```

```
<gap reason="passing truck" extent="5" unit="s"/>
```

```
<u who="#P1">I proposed that <foreign xml:lang="de"> wir können  
<pause dur="PT1S"/> vielleicht </foreign> go to warsaw and <emph>vienna</emph>  
</u>
```

# Participant Description

```
<particDesc>
  <listPerson>
    <person xml:id="P-1234" sex="2" age="mid">
      <p>Female informant, well-educated, born in Shropshire UK, 12 Jan 1950, of
unknown occupation. Speaks French fluently. Socio-Economic status B2. </p>
    </person>
    <person xml:id="P-4332" sex="1">
      <persName>
        <surname>Hancock</surname>
        <forename>Antony</forename>
        <forename>Aloysius</forename>
        <forename>St John</forename>
      </persName>
      <residence notAfter="1959">
        <address>
          <street>Railway Cuttings</street>
          <settlement>East Cheam</settlement>
        </address>
      </residence>
      <occupation>comedian</occupation>
    </person>
  </listPerson>
</particDesc>
```

## Flexibility and standardization of metadata

- Information can be supplied purely in documentary terms, as plain text...
- ... or it can be organized in a structured way, using a rich set of XML descriptors
- The available descriptors can be constrained by means of a customised schema
- Additional descriptors can be added, and integrated by means of the TEI class system

# <scriptStmt> example

```
<sourceDesc>
  <scriptStmt xml:id="CNN12">
    <bibl>
      <author>CNN Network News</author>
      <title>News headlines</title>
      <date when="1991-06-12">12 Jun 91</date>
    </bibl>
  </scriptStmt>
</sourceDesc>
```

## Similarly for recordings...

```
<recordingStmt>
  <recording type="audio" dur="P30M">
    <respStmt>
      <resp>Location recording by</resp>
      <orgName>Sound Services Ltd.</orgName>
    </respStmt>
    <equipment>
      <p>Multiple close microphones mixed down to stereo Digital Audio Tape,
standard play, 44.1 KHz sampling frequency</p>
    </equipment>
    <date>12 Jan 1987</date>
  </recording>
</recordingStmt>
```

# Detailed <recording>

```
<recording type="audio" dur="P10M">
  <equipment>
    <p>Recorded from FM Radio to digital tape</p>
  </equipment>
  <broadcast>
    <bibl>
      <title>Interview on foreign policy</title>
      <author>BBC Radio 5</author>
      <respStmt>
        <resp>interviewer</resp>
        <name>Robin Day</name>
      </respStmt>
      <respStmt>
        <resp>interviewee</resp>
        <name>Margaret Thatcher</name>
      </respStmt>
    </bibl>
  </broadcast>
</recording>
```

## Or maybe just...

```
<recordingStmt>  
  <recording type="audio" dur="P15M" xml:id="rec-3001">  
    <date>14 Feb 2001</date>  
  </recording>  
  <recording type="audio" dur="P15M" xml:id="rec-3002">  
    <date>17 Feb 2001</date>  
  </recording>  
  <recording type="audio" dur="P15M" xml:id="rec-3003">  
    <date>22 Feb 2001</date>  
  </recording>  
</recordingStmt>
```

## ... and for settings

```
<setting xml:id="KDFSE002" n="063505" who="#PS0M6">  
  <name type="place">Lancashire: Morecambe </name>  
  <locale> at home </locale>  
  <activity> watching television </activity>  
</setting>
```

# Analysis and Linguistics

- associating simple analyses and interpretations with text elements
- semantic or syntactic interpretations which an encoder wishes to attach to all or part of a text
- mainly covering linguistic information
- as often in the TEI, you can do the same thing in many ways:
  - using generic <seg> elements with *@type* attributes
  - using the straightforward *canned* analyses described here
  - using the more powerful and general TEI Feature Structures

## Linguistic units

To mark up text for linguistic purposes:

`<s>` (s-unit) contains a sentence-like division of a text.

`<cl>` (clause) represents a grammatical clause.

`<phr>` (phrase) represents a grammatical phrase.

`<w>` (word) represents a grammatical (not necessarily orthographic) word.

`<m>` (morpheme) represents a grammatical morpheme.

`<C>` (character) represents a character.

From the `att.segLike` class, these elements all have `@type` and `@function` attributes

# Example of linguistic markup

## Compare

<u>Like a suck of one of my sweets?</u>

<u>No I don't take sweets from strangers, oh God</u>

with....

# linguistic markup

```

<u who="PS1K5">
  <s n="5963">
    <w type="AV0">Like</w>
    <w type="AT0">a</w>
    <w type="NN1">suck</w>
    <w type="PRF">of</w>
    <w type="CRD">one</w>
    <w type="PRF">of</w>
    <w type="DPS">my</w>
    <w type="NN2">sweets</w> ?</s>
  </u>
<u trans="smooth" who="PS1BY">
  <s n="5964">
    <w type="ITJ">No </w>
    <w type="PNP">I </w>
    <w type="VDB">do</w>
    <w type="XX0">n't </w>
    <w type="VVI">take </w>
    <w type="NN2">sweets </w>
    <w type="PRP">from </w>
    <w type="NN2">strangers</w>
    <c type="PUN">,</c>
    <w type="ITJ">oh </w>
    <w type="NP0">God</w>
  </s>
</u>

```

(from British National Corpus, KSV 5963)

## Mixing analysis with structure

Analytic units often cross structural boundaries. The `<cl>` (clause) elements here cross the verse lines (`<l>`). We can use the `@part` attribute to show how a `<cl>` can be assembled:

```
<div type="stanza">
  <l>
    <cl part="I">Tweedledum and Tweedledee</cl>
  </l>
  <l>
    <cl part="F">Agreed to have a battle;</cl>
  </l>
  <l>
    <cl part="I">For Tweedledum said Tweedledee</cl>
  </l>
  <l>
    <cl part="F">Had spoiled his nice new rattle.</cl>
  </l>
</div>
```

## Or the @next attribute

```
<l>
  <cl next="#c5" xml:id="c3" part="I">For Tweedledum said
  <cl next="#c6" xml:id="c4" part="I">Tweedledee</cl>
</cl>
</l>
<l>
  <cl prev="#c3" xml:id="c5" part="F">
  <cl prev="#c4" xml:id="c6" part="F">Had spoiled his nice new rattle.</cl>
</cl>
</l>
```

## Stand-off interpretation

When inline markup is inappropriate, the `<span>` element can be used to make *ad hoc* remarks about bits of text, linked to by ID. As usual, `<spanGrp>` is available to group assertions together.

```

<sp>
  <speaker>CORNWALL</speaker>
  <ab xml:id="eye_start">Lest it see more, prevent it. Out, vile jelly!</ab>
  <ab>Where is thy lustre now?</ab>
</sp>
<sp>
  <speaker>GLOUCESTER</speaker>
  <ab>All dark and comfortless. Where's my son Edmund?</ab>
  <ab>Edmund, enkindle all the sparks of nature,</ab>
  <ab xml:id="eye_end">To quit this horrid act.</ab>
</sp>
<span from="#eye_start" to="#eye_end">the eye is pulled out</span>

```

## Stand-off interpretation with <interp>

The <interp> element is used to encode an interpretation. The global @ana attribution can point from the text to such an interpretation:

```
<ab n="2Sam_12:14">
  <gap/>by this deed thou hast given great occasion to the enemies of the LORD
  to blaspheme, the child also that is born unto thee shall surely die.
</ab>
<ab n="2Sam_12:15">
  <gap/>And the LORD struck the child that Uriah's wife bare unto David<gap/>
</ab>
<gap/>
<ab n="2Sam_12:18" ana="#infanticide">And it came to pass on the seventh day,
that the child died.</ab>
<!-- elsewhere in document -->
<interp resp="#SAB" xml:id="infanticide">Infanticide: God seems to like
killing children.</interp>
```

The <interpGrp> element is used to group interpretations together.

## Interpretation example (1)

In this example:

- A set of possible interpretations is defined, using `<interp>` elements
- `<seg>` is used to markup distinct portions of a narrative
- `<S>` is used to mark sentences
- the `@ana` attribute links sections or milestones to appropriate interpretation

```
<interpGrp resp="#TMA" type="structuralUnit">  
  <interp xml:id="INTRO">introduction</interp>  
  <interp xml:id="CONFLICT">conflict</interp>  
  <interp xml:id="CLIMAX">climax</interp>  
  <interp xml:id="REVENGE">revenge</interp>  
  <interp xml:id="RECONCIL">reconciliation</interp>  
  <interp xml:id="AFTERM">aftermath</interp>  
</interpGrp>
```

## Interpretation example (2)

```

<p xml:id="PP1">
  <seg xml:id="SS1-SS3" ana="#INTRO">
    <s xml:id="SS1">Sigmund ... was a king in Frankish country. </s>
    <s xml:id="SS2">Sinfiotli was the eldest of his sons. </s>
    <s xml:id="SS3">Borghild, Sigmund's wife, had a brother ... </s>
  </seg>
  <s xml:id="SS4A" ana="#CONFLICT">But Sinfiotli ... wooed the same woman</s>
  <s xml:id="SS4B" ana="#I3">and Sinfiotli killed him over it. </s>
  <seg xml:id="SS5-SS17" ana="#CLIMAX">
    <s xml:id="SS5">And when he came home, ... she was obliged to accept
it. </s>
    <s xml:id="SS6">At the funeral feast Borghild was serving beer. </s>
    <s xml:id="SS17">Sinfiotli drank it off and at once fell dead. </s>
  </seg>
</p>
<anchor xml:id="NIL1" ana="#RECONCIL"/>
<p xml:id="PP2">Sigmund carried him a long way in his arms ... </p>

```

# Linguistic Transcription

When transcribing, some people are more interested in the linguistic values of texts than their physical or semantic contexts.

# Phrase segmentation

```
<s>
  <cl type="finite-declarative" function="independent">
    <phr type="NP" function="subject">It</phr>
    <phr type="VP" function="predicate">
      <phr type="V" function="verb-main">was</phr>
      also
    <phr type="NP" function="predicate-nom.">a crucial year for me</phr>
  </phr>
</cl>
</s>
```

# Words with lemmas and morphemes with types

```
<s xml:lang="la">  
  <w lemma="timeo">timeo</w>  
  <w lemma="danaii">Danaos</w>  
  <w lemma="et">et</w>  
  <w lemma="donum">dona</w>  
  <w lemma="fero">ferentes</w>  
</s>
```

or

```
<w type="adjective">  
  <m type="prefix" baseForm="con">com</m>  
  <m type="root">for t</m>  
  <m type="suffix">able</m>  
</w>
```

# Nested <w>

```
<s>  
  <w>I</w>  
  <w>  
    <w>did</w>  
    <m>n't</m>  
  </w>  
  <w>do</w>  
  <w>it</w>  
  <c>.</c>  
</s>
```

# Word analysis

```
<S>
  <w ana="#AT0">The</w>
  <w ana="#NN1">victim</w>
  <w ana="#POS">'s</w>
  <w ana="#NN2">friends</w>
  <w ana="#VVD">told</w>
  <w ana="#NN2">police</w>
  <w ana="#CJT">that</w>
  <w ana="#NP0">Kruger</w>
  <w ana="#VVD">drove</w>
  <w ana="#PRP">into</w>
  <w ana="#AT0">the</w>
  <w ana="#NN1">quarry</w>
  <w ana="#CJC">and</w>
  <w ana="#AV0">never</w>
  <w ana="#VVD">sur faced</w>
</S>
```

# Interpretation

```
<interpGrp type="POS">  
  <interp xml:id="AT0">Definite article</interp>  
  <interp xml:id="AV0">Adverb</interp>  
  <interp xml:id="CJC">Conjunction</interp>  
  <interp xml:id="CJT">Relative that</interp>  
  <interp xml:id="NN1">Noun singular</interp>  
  <interp xml:id="NN2">Noun plural</interp>  
  <interp xml:id="NP0">Proper noun</interp>  
  <interp xml:id="POS">Genitive marker</interp>  
  <interp xml:id="PRP">Preposition</interp>  
  <interp xml:id="VVD">Verb past tense</interp>  
</interpGrp>
```

## More interpretation

```
<u xml:id="u1">Can I have ten oranges and a kilo of bananas please?</u>
<u xml:id="u2">Yes, anything else?</u>
<u xml:id="u3">No thanks.</u>
<u xml:id="u4">That'll be dollar forty.</u>
<u xml:id="u5">Two dollars</u>
<u xml:id="u6">Sixty, eighty, two dollars. Thank you.</u>
<spanGrp type="transactions">
  <span from="#u1">sale request</span>
  <span from="#u2" to="#u3">sale compliance</span>
  <span from="#u4">sale</span>
  <span from="#u5">purchase</span>
  <span from="#u6">purchase closure</span>
</spanGrp>
```

# British National Corpus

- a snapshot of British English, taken at the end of the 20th century
- 100 million words in approx 4000 different text samples, both spoken (10%) and written (90%)
- synchronic (1990-4), sampled, general purpose corpus
- available under licence; latest edition is BNC-XML (13 March 2007)
- Part-of-speech and lemma tagging
- Uses a variant of TEI XML originally called CDIF

## BNC XML

```

<div level="1" n="1" type="leaflet">
  <head type="MAIN">
    <s n="1">
      <w c5="NN1" hw="factsheet" pos="SUBST">FACTSHEET</w>
      <w c5="DTQ" hw="what" pos="PRON">WHAT</w>
      <w c5="VBZ" hw="be" pos="VERB">IS</w>
      <w c5="NN1" hw="aids" pos="SUBST">AIDS</w>
      <c c5="PUN">?</c>
    </s> </head>
  <p>
    <s n="2">
      <hi rend="bo"> <w c5="NN1" hw="aids" pos="SUBST">AIDS</w>
      <c c5="PUL">(</c>
      <w c5="VVN-AJ0" hw="acquire" pos="VERB">Acquired</w>
      <w c5="AJ0" hw="immune" pos="ADJ">Immune</w>
      <w c5="NN1" hw="deficiency" pos="SUBST">Deficiency</w>
      <w c5="NN1" hw="syndrome" pos="SUBST">Syndrome</w>
      <c c5="PUR">)</c>
    </hi>
    <w c5="VBZ" hw="be" pos="VERB">is</w>
    <w c5="AT0" hw="a" pos="ART">a</w>
    <w c5="NN1" hw="condition" pos="SUBST">condition</w>
  <!-- ... -->
    </s>
  </p>
</div>

```