Identification und Correction of Disfluencies in Spontanous Speech

Jana Besser

DFKI GmbH, Saarbrücken
jbesser@dfki.de

Outline

- Introduction To The Topic
  - Aims
  - Motivation
  - AMI Project
  - Theoretical Background
- Disfluencies
  - General Phenomena
  - Dependent Phenomena
- Annotation
- Summary
- Future Work

Text vs. Speech

- Text
  - Follows grammatical rules
  - Words fully articulated
- Spontaneous Speech
  - Contains **Disfluencies**: linguistic irregularities such as
    - Self-corrections
    - Stutterings and
    - Semantically empty fillers
  - Demand for computational linguistic applications

Aims

- Corpus-based classification of disfluencies in spontaneous speech
  - AMI Meeting Corpus
- Development of an annotation scheme
- Evaluation through multiple annotation
- Preparation for the automatic identification and correction of disfluencies
  - Preprocessing for parsing
  - Disfluency as feature for machine learning
- Current status:
  - Disfluency classification
  - Annotation manual
  - Annotation tool

Motivation

- Current systems (and grammars) based on data from written language

  => Do not need to handle disfluencies
- Take advantage of existing ressources, extend them
  - Parsing spontaneous speech with “text grammar” (e.g., English-Resource-Grammar (ERG), HPSG)
- Speaker analysis
  - Uncertainty or dominance in a conversation
  - Non-nativeness

The AMI Project

- AMI: Augmented Multi-party Interaction
- Consortium of industrial and research institutions
- Technology for augmenting communication between individuals and groups of people

  http://www.amiproject.org
- Recordings of business Meetings (about 100 hours)
  - AMI Meeting Corpus

Method

- Division of transcriptions into dialogue act segments
- Parsing results (ERG): ca. 35% non-parsable segments
- Classification based on
  - Analysis of non-parsable segments.
  - Previous work on the topic
- Annotation

Annotation

- Annotation in XML format
- Annotation manual
  - Introduction to topic
  - Definition of phenomena (+ examples)
  - Notes about “False Friends”
- Annotation on transcript
- Emacs mode with annotation keys

Background

- **Finkler PhD**: Generation of disfluencies in TTS
- **Strassel:** Definition of metadata types (*Simple Metadata Annotation Specification,* Linguistic Data Consortium, 2004)
- **Allwood**: Disfluencies as „speech management“

Disfluencies

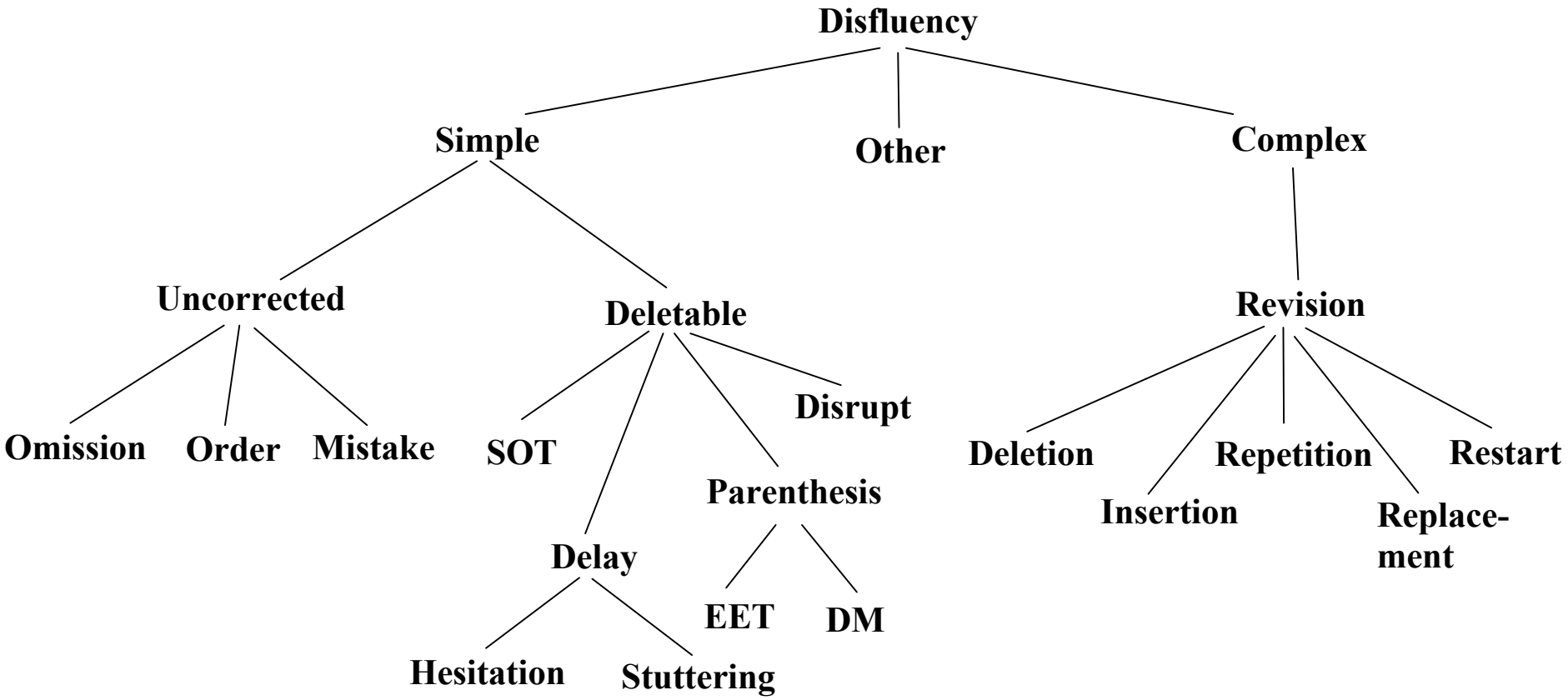
- General structure:

  Reparandum - IP - (Editing Phase) - Reparans

- Three classes of disfluencies
  - Complex
  - Simple
  - Dependent

- Hierarchical structure

Disfluency Hierarchy



Complex Disfluencies
(Correction by speaker)

- Keys:
  - „ |“ – begin/end-of Disfluency
  - „||“ – „interruption point“ between original and correction
  - „Under“ – correction by annotator

- **Deletion** – deletion of speech material (SM)
  But | it‘s really not || it‘s not | functional.
- **Insertion** – Insertion of SM
  | What else it || what else do we want it | to do?
- **Repetition** – Repetition of SM
  Maybe we could draw it up | on the || on the | board.
- **Restart** – Substitution of SM
  So how would we go about | making || getting | rid of our weak points?
- **Replacement** – Partial replacement of SM
  Even if you designed it | in some || in a | way that you know…

XML Format

But it's really not it's not functional.

```
But
    <deletion>
        <reparandum>
            it's really not
        </reparandum>
        <reparans>
            it's not
        </reparans>
    </deletion>
functional.
```

Misc Simple Disfluencies

(Correction by Annotator)

- **Omission – Omission of essential SM**

  And | the | project manager will design a better meeting.

- **Order** – Word order errors

  I don‘t know | what‘s the idea | what the idea is | for

- **Mistake – Grammatical mistake**

  Change | Changing | channels is its main function.

Misc. Simple Disfluencies

- **Hesitation**

  uh, uhm, eh, em, mm
- **Slip Of the Tongue** (SOT)

  Looking at the | tex | technical functions…
- **Stuttering**

  | D | do you really always want to open that thing?
- **Discourse Marker** (DM)

  I mean, you know, anyway, well, so…
- **Explicit Editing Term** (EET)

  | The design of || | or | the point of | putting two…
- **Disruption** – incomplete nonsense speech material

  Or like a

Combinations of Disfluencies

- But then to go back to the to th s something along those things.

```
But then to go back
    <replace>
        <reparandum> to the </reparandum>
        <reparans>
                to
                <sot> th </sot>
                <stutter> s </stutter>
                something
        </reparans>
    </replace>
along those things.
```

Data

(Preliminary Evaluation of Annotations)

- 77% of non-parsable segments contain disfluencies
- Several disfluencies per segment: 287 disfluencies in 184 segments

Discourse Marker:	21%	Stuttering:	3%
Hesitation:	20%	Insertion:	2%
Repetition:	15%	Order:	1%
Disruption:	10%	EET:	1%
Mistake:	7%	Other:	0.7%
Omission:	7%	Deletion:	0.3%
Slip Of the Tongue:	7%	Restart:	0.3%
Replacement:	5%		

Dependent Phenomena

- Phenomena which depend on domain, technology and ressources
  - Slang

    dunno (don't know), sorta (sort of), 'kay (okay)
  - Non-lexicon words
    - Named Entities

      Hi I'm Helena Carter. I'm working for the DFKI.
  - Transcription

    Because of it's stability…
- Semi-dependent: language switches
  - Known expressions: Carpe diem, alia iacta est…
  - Utterances in an unknown foreign language

Conclusion

- Corpus-based disfluency classification
- Advancement and extension of former work (Finkler, Allwood, Strassel)
- Evaluation through comparison of multiple annotations
- Next tasks:
  - Automatic correction of disfluencies
  - Considering use(fulness) of audio/video under investigation