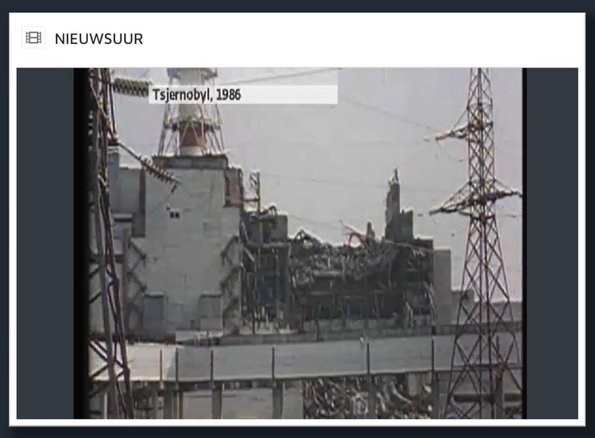


Nieuwsuur. NOS. 17 March, 2011.
https://mediasuite.clariah.nl/tool/resource-viewer?id=210160813011029793 1&cid=daan-catalogue-aggr

Mary-Joy van der Deure
Jasmijn van Gorp
Utrecht University

A 'Local' Method: Analysing Automatic Speech Recognition Files of Dutch TV News

NOS Journaal. "13 uur Journaal." NOS. 26 april, 2006.
https://mediasuite.clariah.nl/tool/resource-viewer?id=210160804093130253 1&cid=daan-catalogue-aggr

Introduction

- In recent decades, broadcast archives have opened up their collections with **Automatic Speech Recognition (ASR).**
- For archives, ASR is predominantly used as a form of metadata enrichment to increase the chances of retrieval: ASR translates spoken dialogues into text and **enriches the metadata**, therefore increasing the chances of discovery (e.g. Ordelman and van Hessen 2018).
- We move beyond the use of these ASR-data as mere transcripts, and approach this data **'with care'.** By doing so, we maintain and respect its characteristics and shift perspectives to the methodological opportunities and challenges of this data.

Case Study

- Dutch televised news of the 1986 **Chernobyl** Nuclear Disaster
- Archived in the ***Netherlands Institute for Sound and Vision***

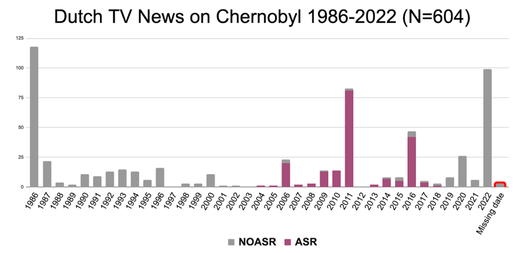


Figure 1: Overview of TV News broadcasts on Chernobyl between 1986 and 2022 without ASR (grey) and with ASR (purple) (N=604)

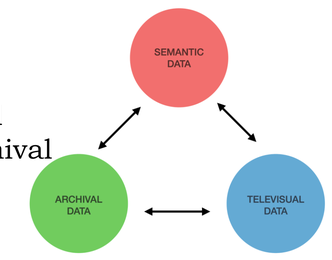
A Local Method

- In line with the field of Critical Data Studies, we consider ASR data as **'locally inscribed'** (Loukissas 2019).
- As with all data, they are an assembly of social, technological and spatial practices that leave their traces. Research with attention for these local inscriptions goes beyond data as "raw input" and instead connects "data back **to the context in which they were produced"** (D'Ignazio and Klein 2020, 152).
- This means we consider ASR to be **an interplay between semantic, televisual and archival data.**

- We formulate and apply a **'local method'** as defined by Loukissas (2019).
- He describes six local features that shape data, of which we selected four. These features help us identify the archival and televisual traces in the data. These features are:
  - **Rituals:**
    We consider both journalistic as well as archival rituals in the form of decisions that are being made by practitioners.
  - **Errors:**
    We do not clear our data, but we read the ASR errors as traces of the data's production.
  - **Absences:**
    By also considering what is missing, we gain insight into what the archive is and what it is not preserving, what is receiving ASR priority and what is not present at all due to journalistic decisions.
  - **Schemata:**
    We also consider the ways the ASR is brought to us, which includes the way this is recorded and inscripted.
- Lastly, we share and discuss our findings with **practitioners** in the field in order to better understand our findings and to grasp ASR's local context.
- This analysis is based on manual annotations of ASR-files and metadata, combined with the videos themselves as present in the CLARIAH Media Suite.
-

ASR as Layered Data

- **ASR as semantic data:** The layer in which ASR is predominantly studied. This entails approaching ASR-data as linguistic transcripts of the video or radio broadcast (De Quintana and Carpio 2021).
- **ASR as televisual data:** This means that traces of the ASR's production and broadcasting context are present in the form of characteristics such as framing, re-use and flow (Williams 1974).
- **ASR as archival data:** This means that the ASR-data is embedded with traces of the choices made by archivists, as well as the technology used. What archivists keep and what they delete depends on an archive's policy and archival practices that in turn take place in a specific historical context (Ketelaar 2001).



Error Typology

- **Substitution:** The ASR has replaced a word by another word.
- **Insertion:** The ASR has inserted a word that was not spoken at all.
- **Deletion:** The ASR has left words completely out of the transcript.

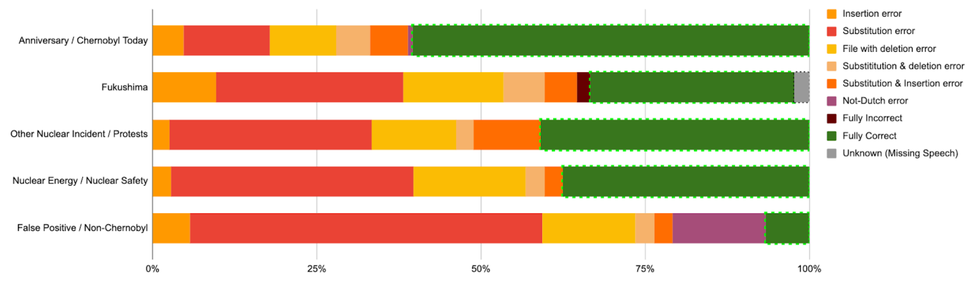


Figure 2: Error typology per topic.

Dissecting the Data: Topics

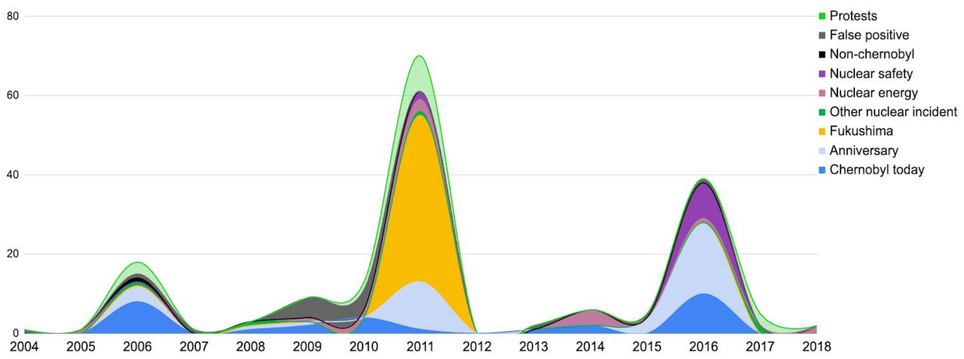
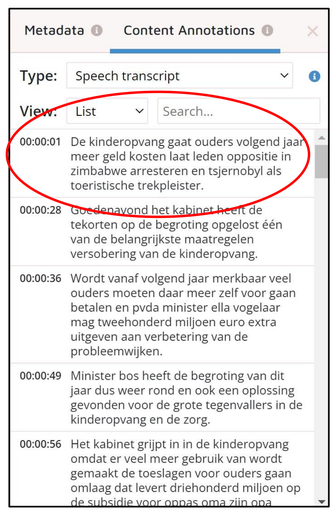


Figure 3: Stream diagram of topics through the ASR-years 2004-2018 (N=171)

Headlines

- While appearing as a strange sentence, knowledge of the ASR schemata tells us this is actually a **televisual headline**.
- This is visible through the **timestamp**, as well as the **grammar** of the sentence, as the ASR is not able to include comma's in summarising sentences like this.



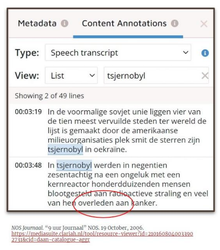
['**00:00:01 -** Childcare is going to cost parents more money next year has members opposition in zimbabwe arrested and chernobyl as tourist attraction.']

Repetition

19-10-2006 - 9am
[**00:06:48** - Hundred of thousands of people were exposed to radioactive radiation in chernobyl in nineteen eighty six after an accident with a nuclear reactor and many of them died of cancer.']

19-10-2006 - 10am
[**00:07:55** - Hundred of thousands of people were exposed to radioactive radiation in chernobyl in nineteen eighty six after an accident with a nuclear reactor and many of them now suffer from cancer.']

19-10-2006 - 1pm
[**00:10:48** - Hundred of thousands of people were exposed to radioactive radiation in chernobyl in nineteen eighty six after an accident with a nuclear reactor and many of them now have cancer.']



NOS Journaal. "9 uur Journaal." NOS. 19 October, 2006.
https://mediasuite.clariah.nl/tool/resource-viewer/id-2016080402031002 77?filecid-daan-catalogue-aggr

- Inclusion of televisual characteristics, like repetition of news items.
- This results in insight into script adjustments throughout the day.
- **Journalistic rituals** like this thus leave their traces in the ASR.

Bibliography

- De Quintana, Eugenio López and Antonio Léon Carpio. 2021. "Artifical intelligence for a role change in television archives: The Artresmedia-Etiqmedia experience." *Journal of Digital Media Management* 10, no. 2 (December): 177-187.
- D'Ignazio, Catherine and Lauren Klein. 2020. *Data Feminism.* London: The MIT Press.
- Ketelaar, Eric. 2001. 'Tacit Narratives: The Meanings of Archives'. *Archival Science* 1 (2): 131–41. https://doi.org/10.1007/BF02435644.
- Loukissas, Yanni Alexander. 2019. *All data are local: thinking critically in a data-driven society.* London: The MIT Press. https://doi.org/10.7551/mitpress/11543.001.0001.
- Ordelman, Roeland, Liliana Melgar, Jasmijn van Gorp and Julia Noordegraaf. 2019. "Media Suite: Unlocking Audiovisual Archives for Mixed Media Scholarly Research." *Selected papers from the CLARIN Annual Conference, Pisa, 8-10 October 2018,* 133-143. Linköping: Linköping University Electronic Press.
- Williams, Raymond. 1974. *Television: Technology and Cultural Form.* Oxfordshire: Taylor & Francis.

Utrecht University · CLICK NL · BEELD & GELUID · MAKE MEDIA GREAT AGAIN · Hogeschool van Amsterdam · MEDIA SUITE