A multidisciplinary approach to the use of technology in research: the case of interview data

Louise Corti
Christoph Draxler

Parthenos workshop for
CEE Countries
7-9 October 2019,
Sofia

Welcome

Thank you for joining us here in Sofia!

Part of Parthenos workshop for CEE Countries

Thank to the great organising team!

Who are we?

A multidisciplinary group of European scholars

- tools and data professionals
- fields - speech technology, social sciences, human computer interaction, oral history and linguistics

Interested in strengthening the position of **interview data** in Digital Humanities

Arjan van Hessen
Stef Scagliola
Silvia Calamai
Henk van den Heuvel

Christoph Draxler
Louise Corti
Jeannine Beeken
Norah Karrouche

Thanks to our Oral Hisory.eu colleagues

- Stefania Scagliola, University of Luxembourg
- Silvia Calamai, University of Siena
- Norah Karrouche, Erasmus University Rotterdam
- Jeannine Beeken, University of Essex
- Arjan van Hessen, University of Twente
- Henk van den Heuvel, Radboud University
- Maureen Haaker, University of Essex
- Max Broekhuizen, Erasmus University Rotterdam

Aims of our multidiscplinary work

- Explore the applicability and usefulness of existing infrastructure & tools for non-digital humanities approaches

- Elucidate why '**language & linguistic' tools are not typically used** by social science or humanities scholars

- Identify the barriers to using such methods and tools
  - paradigms, methods, jargon, technical ability, tool familiarity

- Explore **how** these can be utilised by scholars using interview data

An open mind...expanding one's toolbox

- Consider **preparing - reading - listening - viewing**
- Features of audiovisual and textual data that users may not have previously considered
- We believe that open-source tools can offer benefits to preparing 'data' and to interpreting them
- Workshop feedback very useful to gather uEx

Background and user gatherings

- Stream of activity around the exploitation of techniques and tools for working with oral history (OH) data
  - Oralhistory.eu site: Oxford, Arezzo, Munich and Utrecht

- Explore the diversity of scholarly practices across disciplines who use interview data sources in their daily work: digital humanities, linguistics, oral history and traditional social science

- Previous 4 workshops - tools specialists, data stewards and scholars

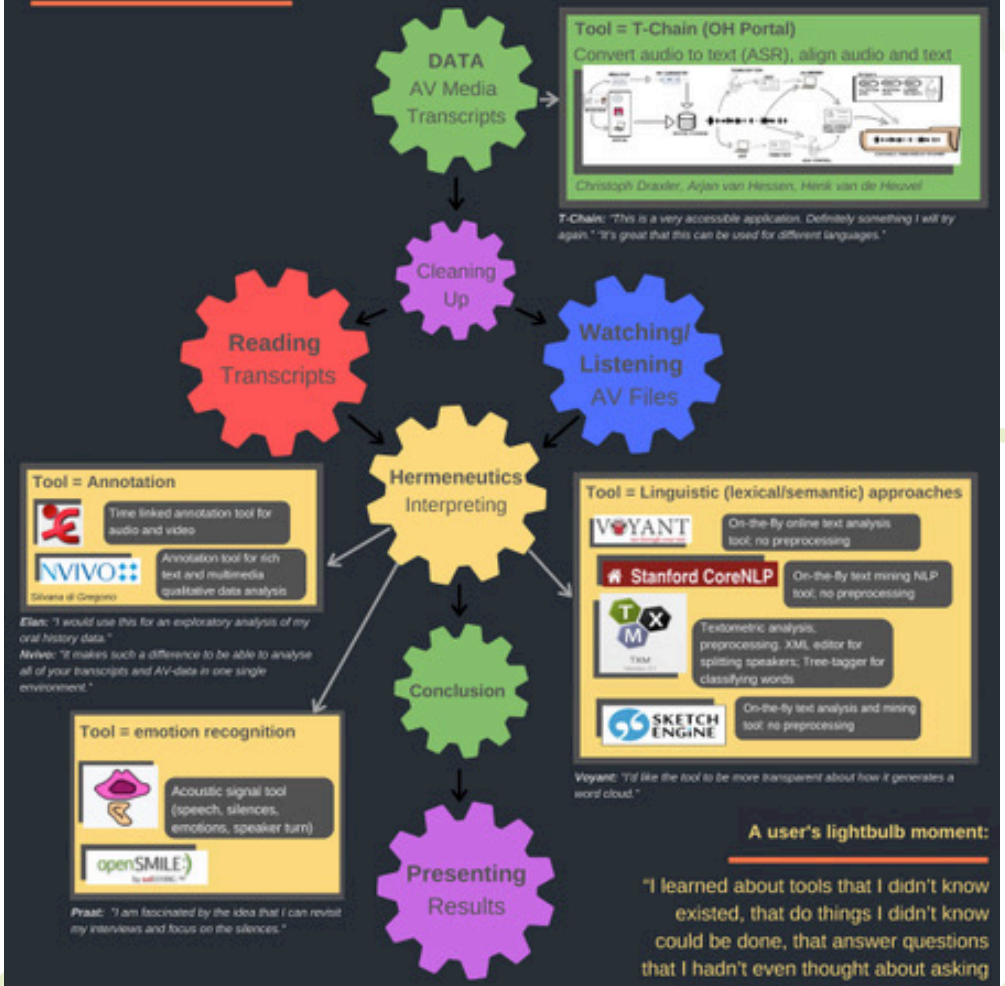
Types of tools our workshops have explored

- Transcribing and aligning oral sources:
  - TChain
- Annotating text - using pre processing or not:
  - ELAN, NVivo and TXM
- Applying linguistic tools to oral history data
  - Voyant, SketchEngine, StanfordCoreNLP,
- Using emotion recognition tools
  - OpenSmile, Praat

Oral History & Technology

We are grateful to:

Since 2016 a multidisciplinary group of speech technologists, social scientists, linguists and oral historians have come together to explore the integration of digital tools in the existing workflows of scholars who work with oral history and interview data



A user's lightbulb moment:

"I learned about tools that I didn't know existed, that do things I didn't know could be done, that answer questions that I hadn't even thought about asking and that I had no awareness that I might be interested in."

Organizers: Arjan van Hessen (Utrecht University / University of Twente), Louise Corti (UK Data Service), Stef Scagliola (University of Luxembourg), Silvia Calamai (University of Siena), Christoph Draxler (Ludwig-Maximilians-Universität München), Jeannine Beeken (UK Data Service), Norah Karrouche (Vrije Universiteit Amsterdam), Max Broekhuizen (Vrije Universiteit Amsterdam), Maureen Haaker (UK Data Service), Khiet Truong (University of Twente)
See: http://oralhistory.eu

Structure of our workshop

Part 1: Lecture

Introducing different scholarly approaches when working with interview data as a primary or secondary data source: distinct traditions and differences in analytic practices and use of tools by scholars across the disciplines

Automated speech recognition, annotation, text analysis and emotion recognition tools are open to wider exploitation

Presentation of TChain tool

Your experiences

Structure of our workshop

Part II Demo and hands-on

The Transcription Chain (TChain) - a tool to convert audio-visual material into a suitable format, use automatic speech recognition (ASR), correct the ASR results, and download them

Discussion of potential of the T-Chain and whether a call for other languages to be integrated into the T-Chain

Emotion recognition video

Exercise & group discussion of your own scholarly work flows and tools when/if analysing interview text

Information - shared drive

The workshop materials

Tiny url: http://tiny.cc/sg93dz

A multidisciplinary approach to the use of technology in research: the case of interview data

Louise Corti
and colleagues

Parthenos workshop for
CEE Countries
7-9 October 2019,
Sofia

The challenge - crossing disciplines

Working with interview data as a primary or secondary data source

- Vast mounts of interview data waiting to be analysed….
- New and often unexpected ways of confronting the raw data..'new questions of old data'
- But when it comes to techniques and tools, researchers are siloed..choosing the familiar over novel
- Observance of Tradition

Different scholarly approaches

- Distinct traditions and differences in analytic practices and use of tools by scholars
  - Automated Speech Recognition
  - transcription
  - annotation
  - text analysis
  - emotion recognition tools

- Potential for digital humanities scholars, historians and social scientists to use speech technology, descriptive and analytical tools for language & linguistic analysis of interview data

When, where and how?

- These approaches can support different phases of the research process, from data preparation to analysis and presentation
- But the ‘whole tool set’ will be unfamiliar
- And, tools are very disconnected….
- which and can thwart the needs of a researcher’s analytic journey
- and hinder new explorations!

Landscape

What do social scientists do with interviews?

Dimensions of an interview

Focus	Research question topics
The Interview	Identity (e.g. race, ethnicity, gender, age, etc.), experience and ”truth”
Recording of the interview	Power, performance
Transcription of the interview	Narratives, memory
Interpretation of the interview	Subjectivity and intersubjectivity; emotion; cultural representation

Abrams (2010)

	Research questions	Approach
Content analysis	What is the amount and nature of attention/reporting on a subject in a specific media?	Largely quantitative
Conversation analysis	How is social order accomplished in and through interaction in everyday life?	Ethnography and conversation analysis
Discourse analysis	What is the discourse doing? Why some meanings become priviledge or taken for granted and others become marginalized?	Study of well-established meanings/ideas around a topic which shape how we talk about it

Adapted from Bryman (2016)

	Research questions	Approach
Narrative analysis	How do people make sense of what happened and to what effect?	Interpretation not dependent on veracity of the account • narrative thematic analysis - focus on content and sequence - what is said/not said • structural analysis - focus on language and linguistic practices - how is the story told • Interactional context - focus of co-construction of narratives - why, when, to whom, in what context
Thematic analysis	What are recurrent ideas and topics in the data?	Look at patterns across interviews. Veracity of the account important. Can be combined with narrative analysis

Annotation

Texts, images, audio, video - approaches to annotation differ across disciplines

- Linguistics
- Behavioral sciences, communication studies
- Media studies (film, performance, dance, television)
- Social sciences
- Digital humanities

Different tools allow for these perspectives

Landscape

What do oral historians do with interviews?

Slides by Norah Karrouche and Max Broekhuizen

Oral history as interviewing practice

- Oral history methods mostly deal with oral history as a **practice** and as an **approach** to history (minority groups, voice)
- Focus on data gathering and publishing the oral histories
- Weak tradition of reflecting on analytical frames

Key questions Oral History

What can this person tell me about the past that I cannot find in the archive? (narrative is treated as a factual source)

After the linguistic turn (mid 1980s)

Narrative is treated as an interpretation of a memorized personal past

How does this person perceive theirself looking back on the past?

How do they attribute meaning to their experiences in the past?

Oral history as testimony and as identity

Stories of events, people (and places)

- Interviews **complement** archival sources

- Life stories
  - Interviews gathered in a particular style
  - Historians (oral historians, biographic approaches, ethnographic approaches)
  - Construction of **identity**, intersubjectivity

- Testimonies
  - Historians (oral history, social history, cultural history)
  - Mediation of memory of **events**

Ideal-typical analytical approaches

Holistic

- One or more interviews
- Interpretation of a completed sequence of events or a **whole** life story
- *Life story of an immigra*

Categorical

- Parts of one or more interviews
- Identifying and interpreting **specific parts** of one or more interview
- *g home in life story of an ant*

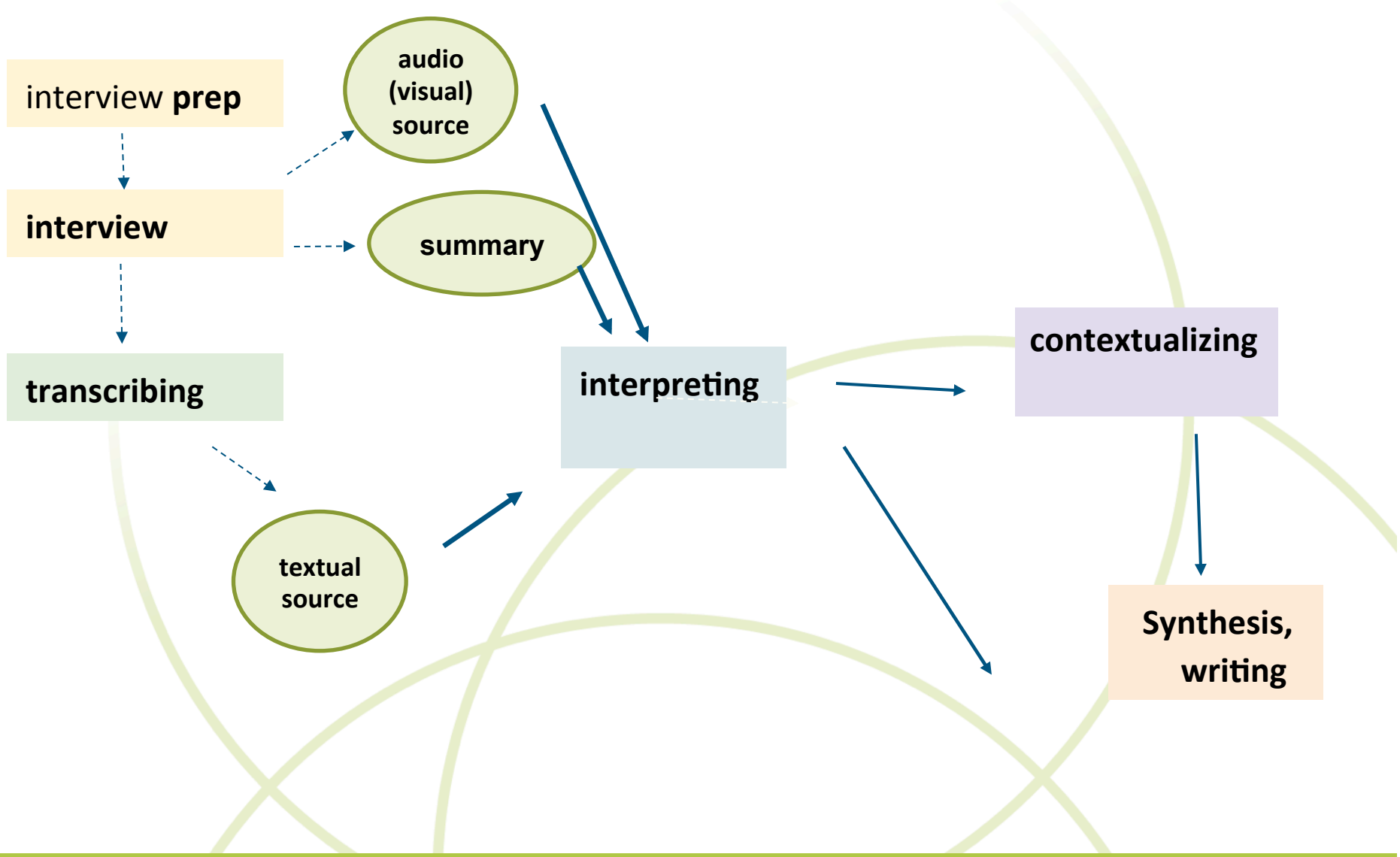
In practice, we use combinations of these approaches ('bricolage')

Thematic

- (Parts of) one or more i
- Content: **what** is being said?
- *The experience of leaving home*

- of) one or more interviews
- Form: **how** is it being said?
- *The way in which the experience of leaving home is narrated (e.g. traumatic, heroic etc.)*

What do oral historians do with interviews?



Landscape

What computational linguists do with spoken corpora?

Slides by Jeannine Beeken

What computational linguists do with spoken corpora?

- **Formulating**
  - research **questions**
  - **hypotheses** to be confirmed/disconfirmed
- **Exploring**
  - and selecting **data** on certain criteria (representativity, language, availability, format, quality, etc.)
  - Testing, selecting and evaluating **tools** for analysis
- **Pre-processing** (when needed)
  - data **transformation** (e.g. speech to text, transcription), cleaning (e.g. errors correction), conversion (one format to another)
  - data **enrichment** (identification of speakers, roles, characteristics - age, gender, language -, uncertain and missing elements, etc.)
  - data **structuring/grouping** (chronology, language, typology of spoken data, etc.)

What computational linguists do with spoken corpora?

- **Analysing**
  - applying methods/tools for **part of speech tagging**, **lemmatization**, frequency **counts**, **comparing**, **clustering**, **parsing**, **disambiguation**, computing **semantic similarities**, recognition of **named entities**, **sentiment** analysis, etc.

- **Interpreting**
  - analysis results taking also into account the **context** (geo-historical, social, political, linguistic, methodological - in collecting and processing the data, etc.)

- **Evaluating**
  - **added value**, **limitations**, **biais**, possibility of **generalisation**, range of **errors**, formulating further **questions/hypotheses** to be studied, etc.

Verbal vs. non-verbal

- Verbal: spoken (sounds, speech) vs. written (text, silent)
- Non-verbal: signs, body language and symbols (silent)
- Excerpt from a silent film https://www.youtube.com/watch?v=4QrR_1NW_w4 (1.10min)

Verbal communication: oral vs. written

- Speech vs. silence
- Spelling: punctuation marks, capital letters, apostrophe, brackets
  - Your first/family name with/without capital
  - Paris Jackson, Orlando Bloom, mother(-)of(-)P/pearl
  - Homophones: seas, sees, seize; friar, fryer; nun, none; grease, Greece
  - Silent letters (speech): knead, kneed (need); knows (noes, nose)
  - Homographs: minute, lead, object, tear
  - Homophone + homograph: arms, ball, spring, duck, watch, ring

Key challenges for linguistic tools

Separate/Disambiguate

- Tokenizers (characters, words etc)
- sentence splitters
- Part of Speech (POS) taggers (nouns, verbs etc.)
- Parsers (grammatical structure)
- Named Entity Recognition (NER)
- Word Sense Disambiguation (WSD)

Key challenges for linguistic tools

Cluster/Group

- Spell checkers
- Lemmatizers (walk, walks, walking, walked)
- Synonyms
- Multiword Expressions (MWE)
  - collocations, idioms, coreferencing

Key challenges for linguistic tools

Control/Reduce

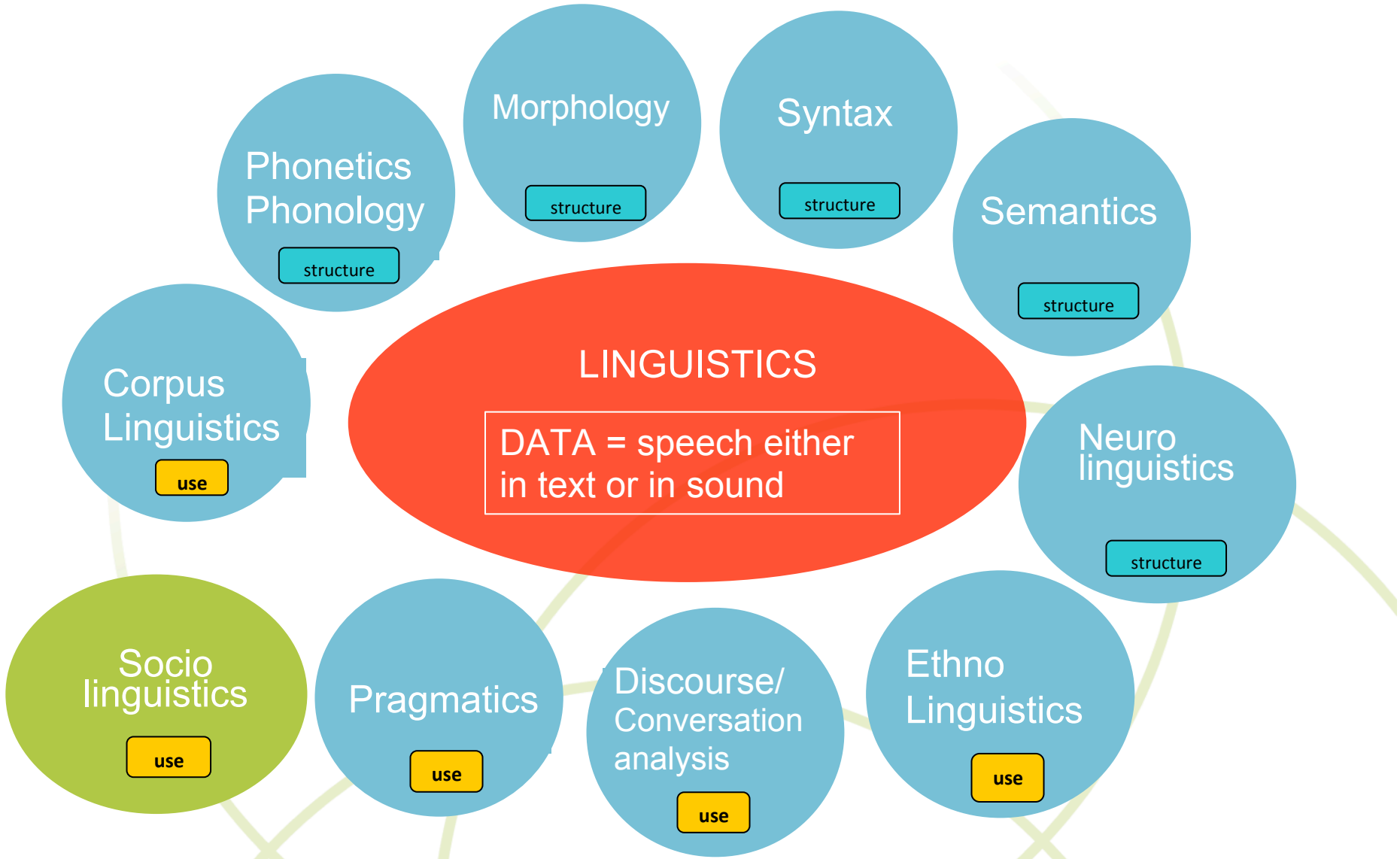
- Stopwords; whitelists and blacklist
- Keyword extraction (with, without a CV)
- Summarizers

Landscape

What do sociolinguists do with spoken corpora?

Slides by Silvia Calamai

Sociolinguists meets qualitative research?


Morphology
structure
Syntax
structure
Phonetics
Phonology
structure
Semantics
structure
Corpus
Linguistics
use
LINGUISTICS
DATA = speech either
in text or in sound
Neuro
linguistics
structure
Socio
linguistics
use
Pragmatics
use
Discourse/
Conversation
analysis
use
Ethno
Linguistics
use

Key questions Sociolinguists

WHO IS SPEAKING ?

- How long? How do they take turns?
  - concepts: overlap/interruption; dialogic repetition; free indirect discourse; polyphonic monologues…)

WHAT AND HOW?

- What are the topics of the interview, semantics, frequency of words
- What are the styles? What does style tell us about the speakers?
- Why do the speakers speak the way they speak?
- Why do they change style and when?
  - (concepts: accommodation, entrainment, linguistic repertoire)

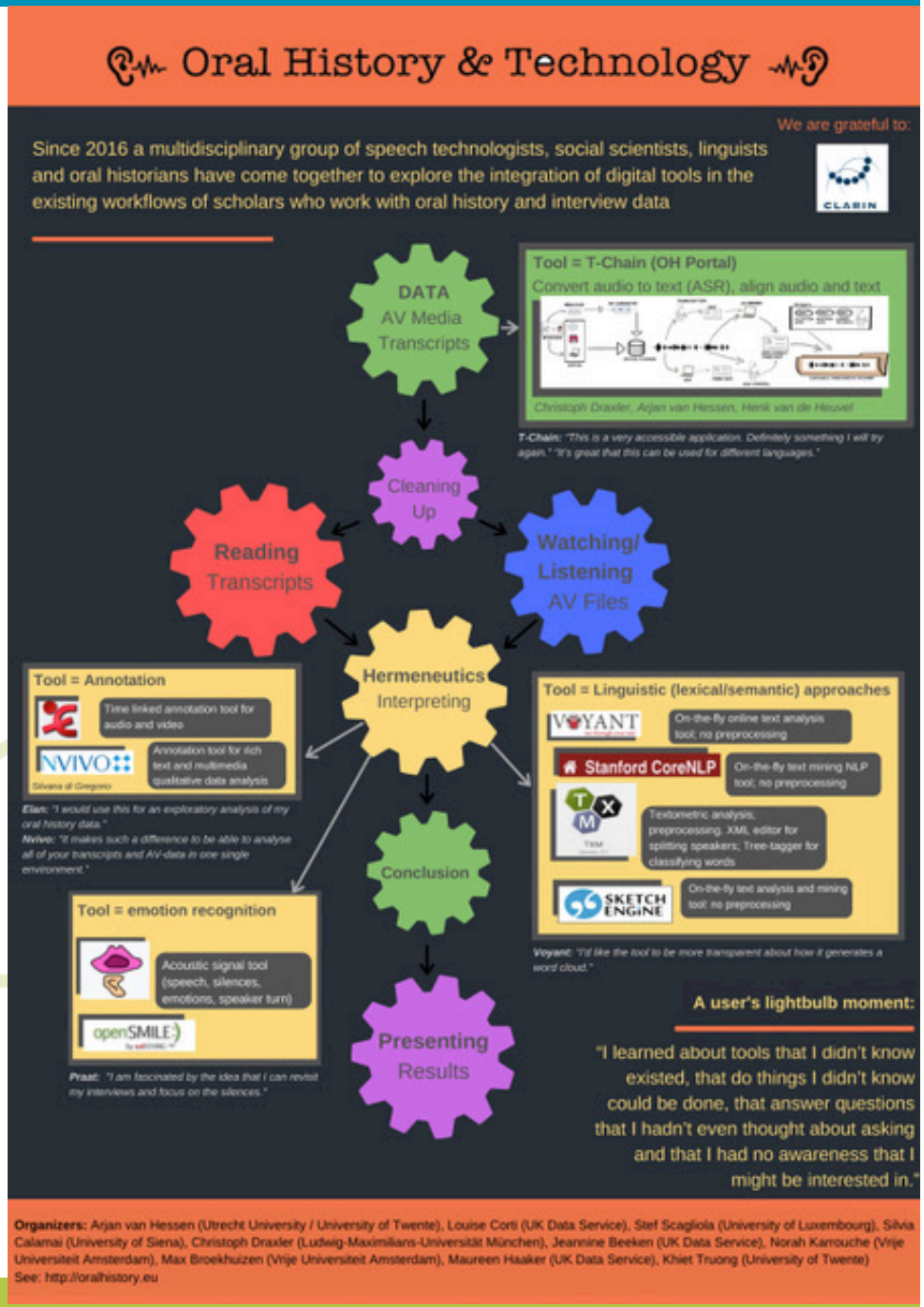
TOOLS for the jobs

Discover

Prepare

Analyse

Report



Discovery: In depth interview data

- Hard to find publicly available FAIR data sources
- That are Findable /discoverable
- Often inAccessible: GDPR, consent, copyright
- Rarely are they Interoperable: formats/metadata
- Reusable? Some have good provenance information

Available qualitative collections

CESSDA archives: https://www.cessda.eu/

From academic research studies

Arising from donations or Research Data Policy

- UK Data Service - self deposit / curated collections
- DANS - self deposit / curated collections
- Others

From local cultural enterprise

- Museums and libraries (National Life Story Collection)
- Local records offices

How NVivo supports analysis

Organizing

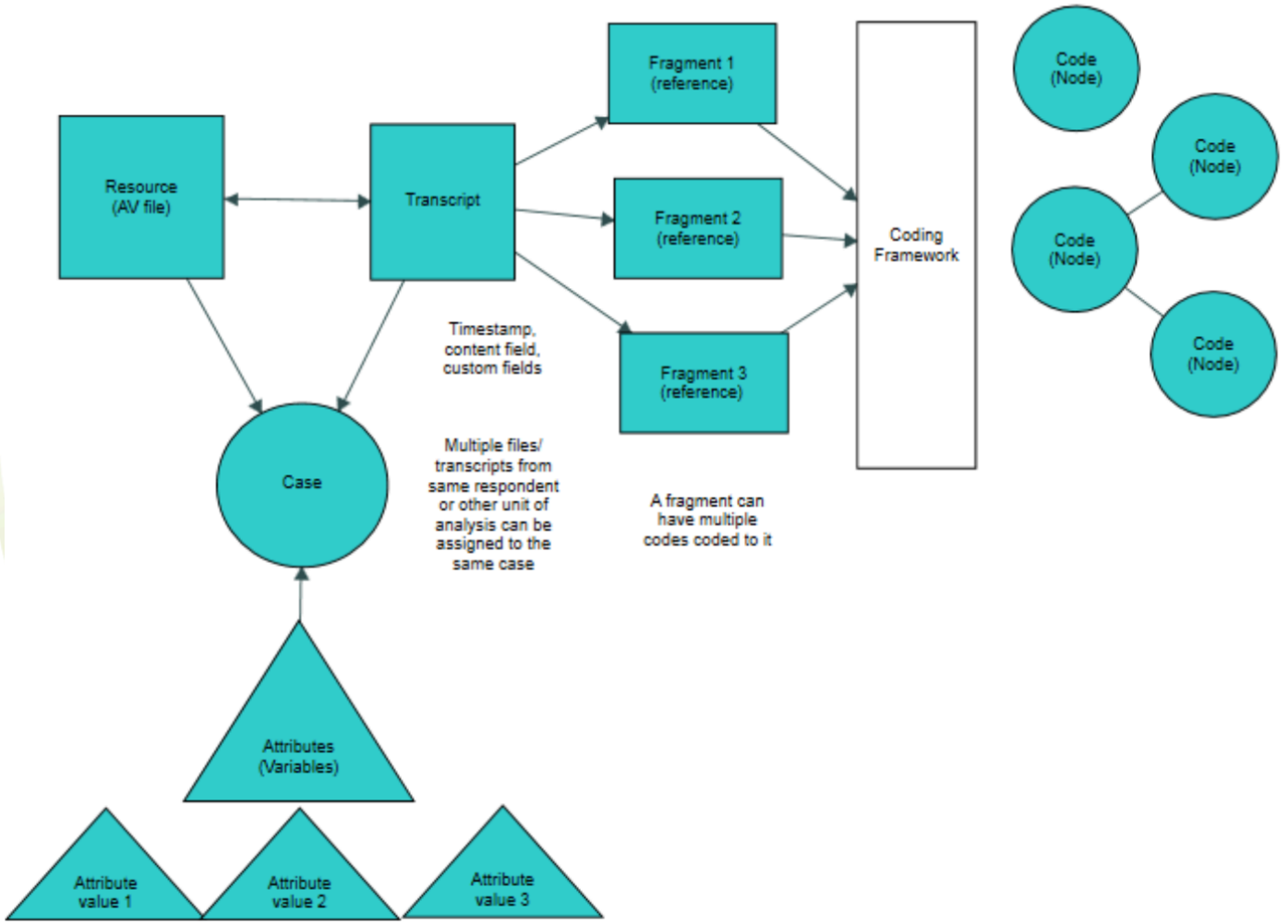
Reflecting

Exploring

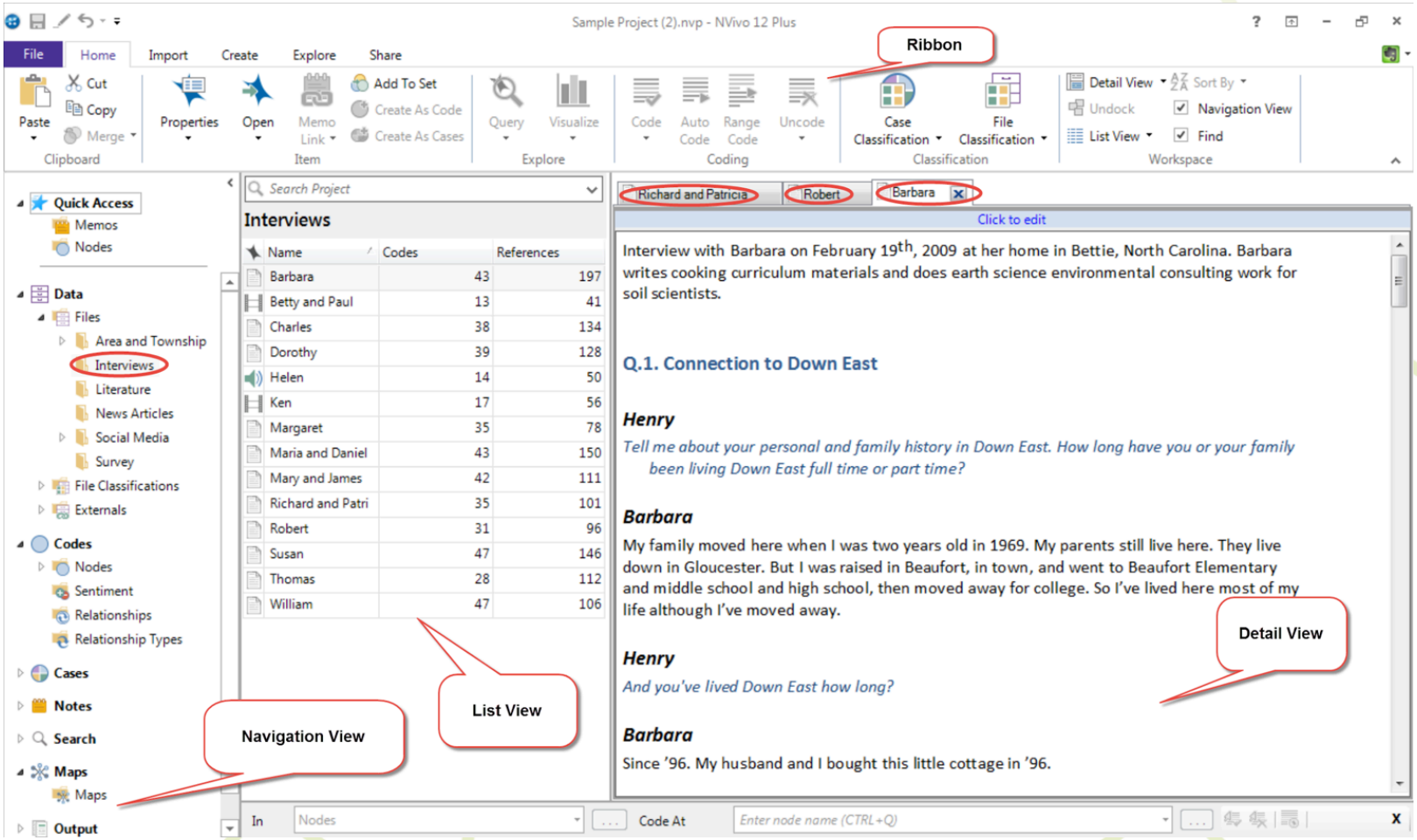
Integrating

Thanks to Silvana di Gregorio for slide input

NVivo Structure



The NVivo Workspace - Windows



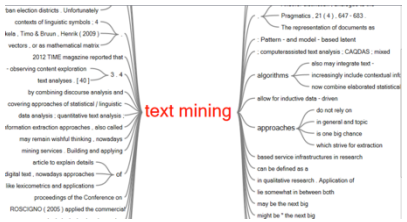
NVivo Analysis Tools

Translation of social science manual methods

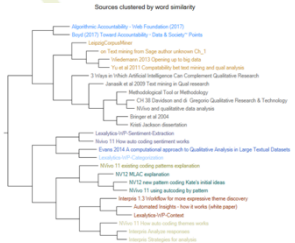
- Coding with stripes (within datum)
- Code retrieval (across data)
- Catalogue system
- Text search
- Query coding
- Cases
- Variables
- Models/ maps
- Comparison diagram
- Sociogram
- Explore diagram

Basic corpus linguistic tools

Word frequency



Word tree - KWIC



Clustering by words or coding

Artificial Intelligence tools

- Machine Learning autocoding
- Autocoding by theme
- Autocoding by sentiment

ELAN (EUDICO Linguistic Annotator)

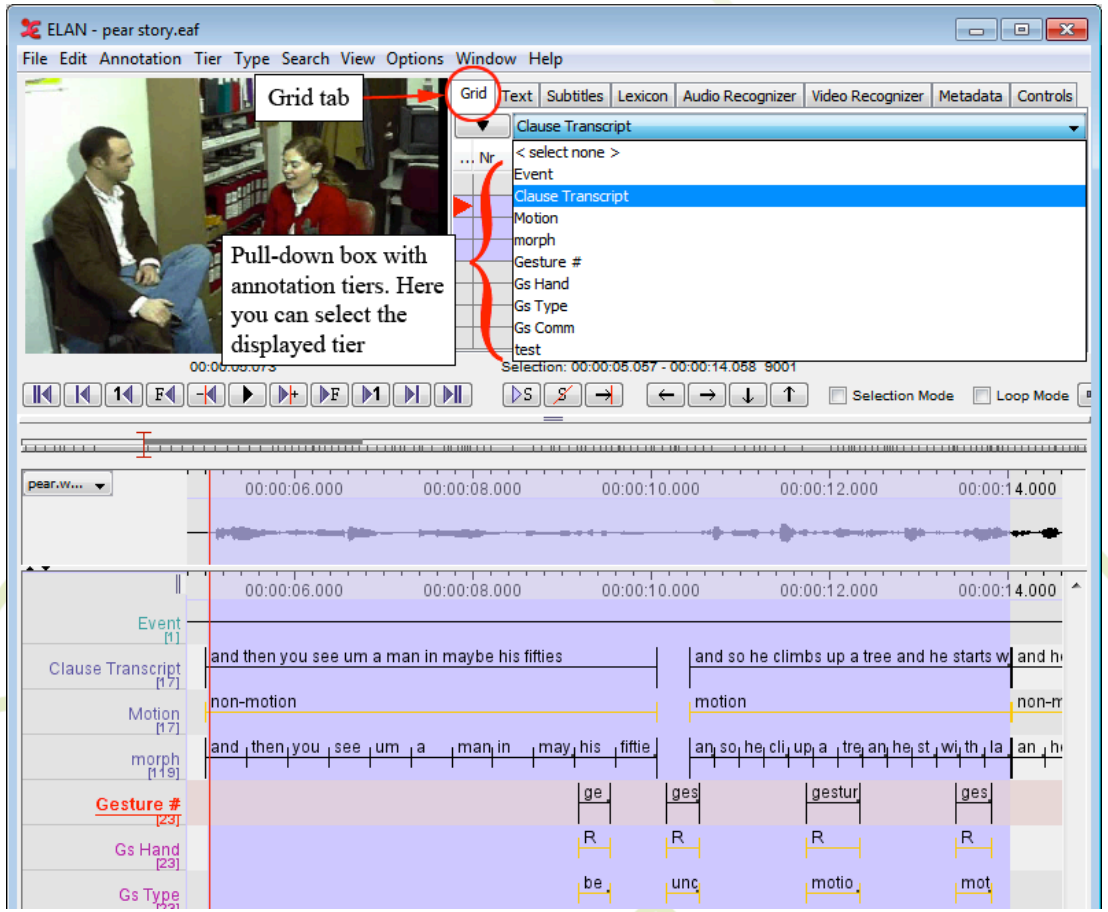
- Open source desktop tool for manual (computer-assisted) annotation (and semi-automatic enrichment) of multimedia
- Developed at the Max Planck Institute for Psycholinguistics, Nijmegen (The Netherlands)
- **Multi-tier, multi-speaker, time-linked annotation of audio and video recordings**
- Produces xml files (.eaf)

Thanks to Lilian Melgar for slide input

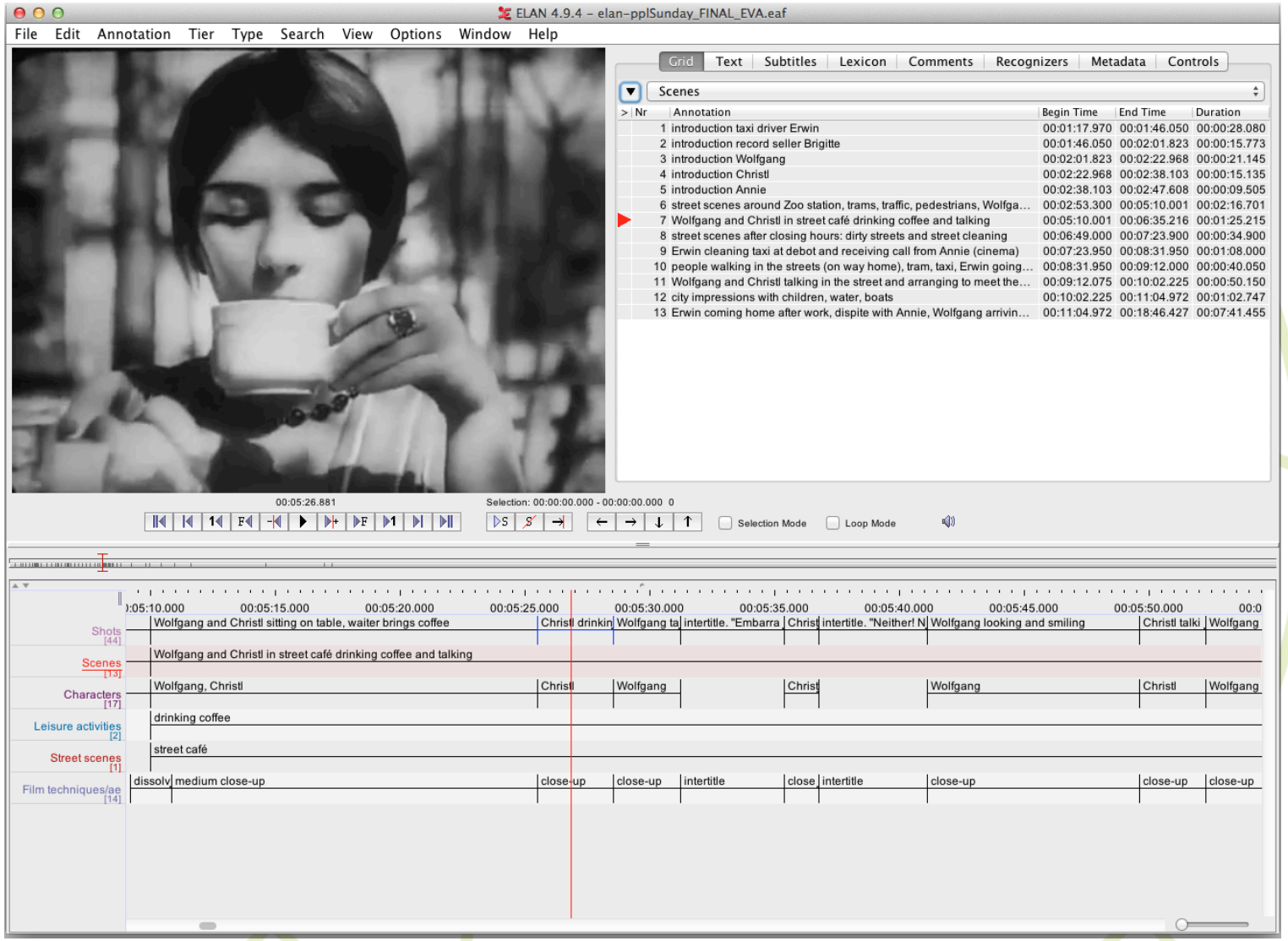
Use of Tiers

- ***"A tier is a set of annotations that share the same characteristics"***

- Tiers are "facets" (same type of annotation, or "event")
- They represent different analytical aspects
- Tiers contain annotations
- They can be grouped hierarchically

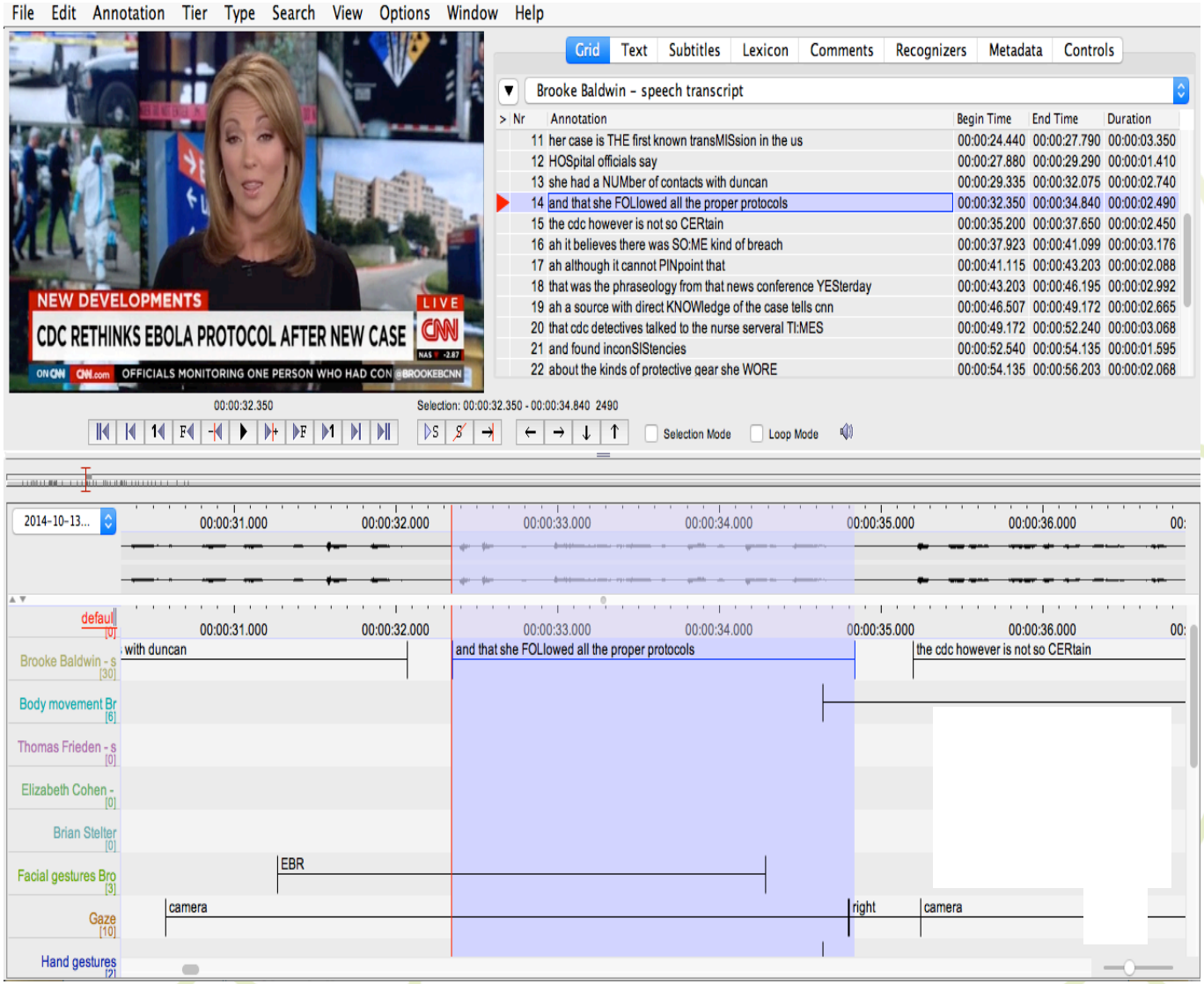


Tiered approach - 1

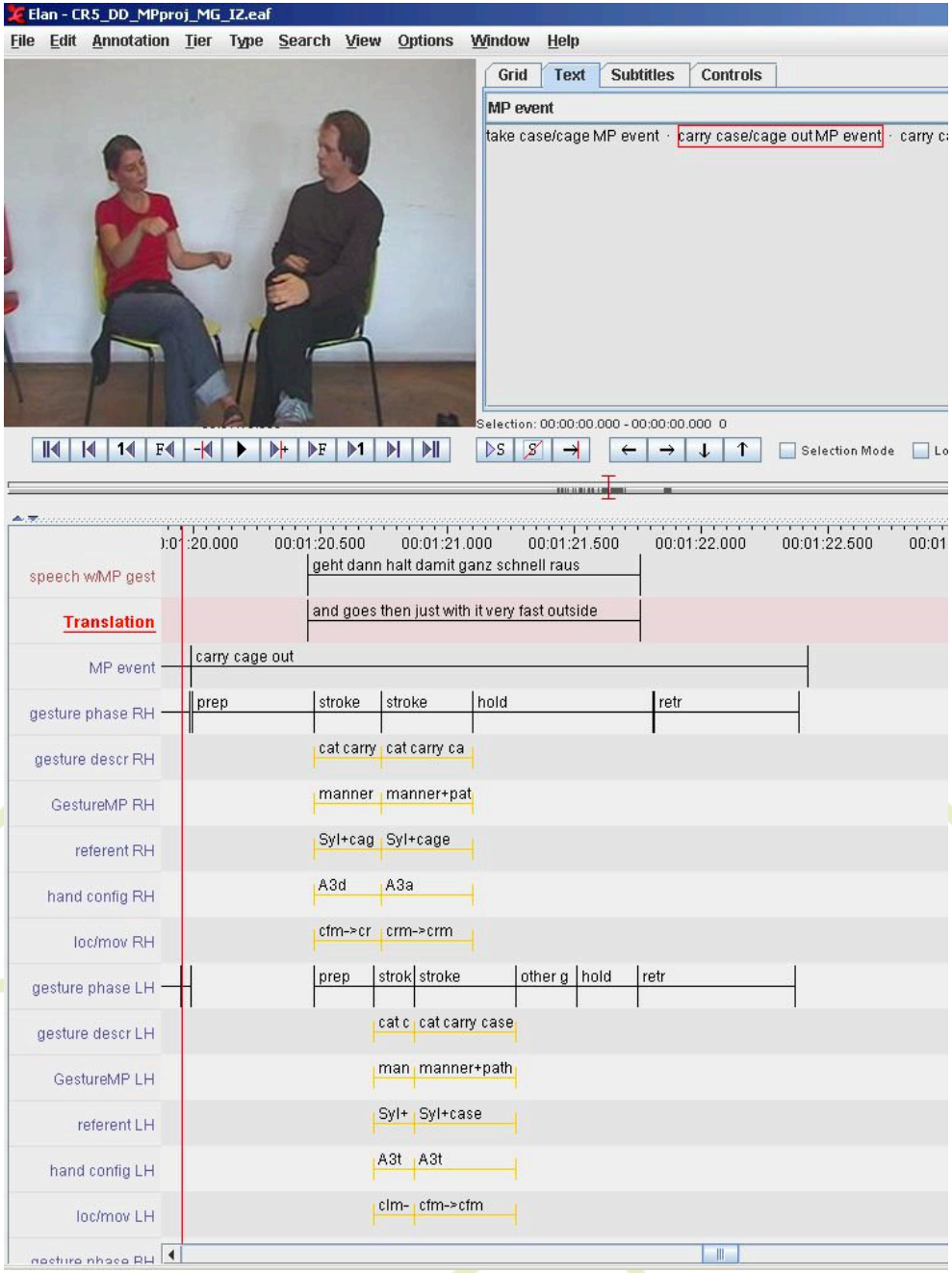


More details in: Sloetjes (2012); Melgar et al., 2017

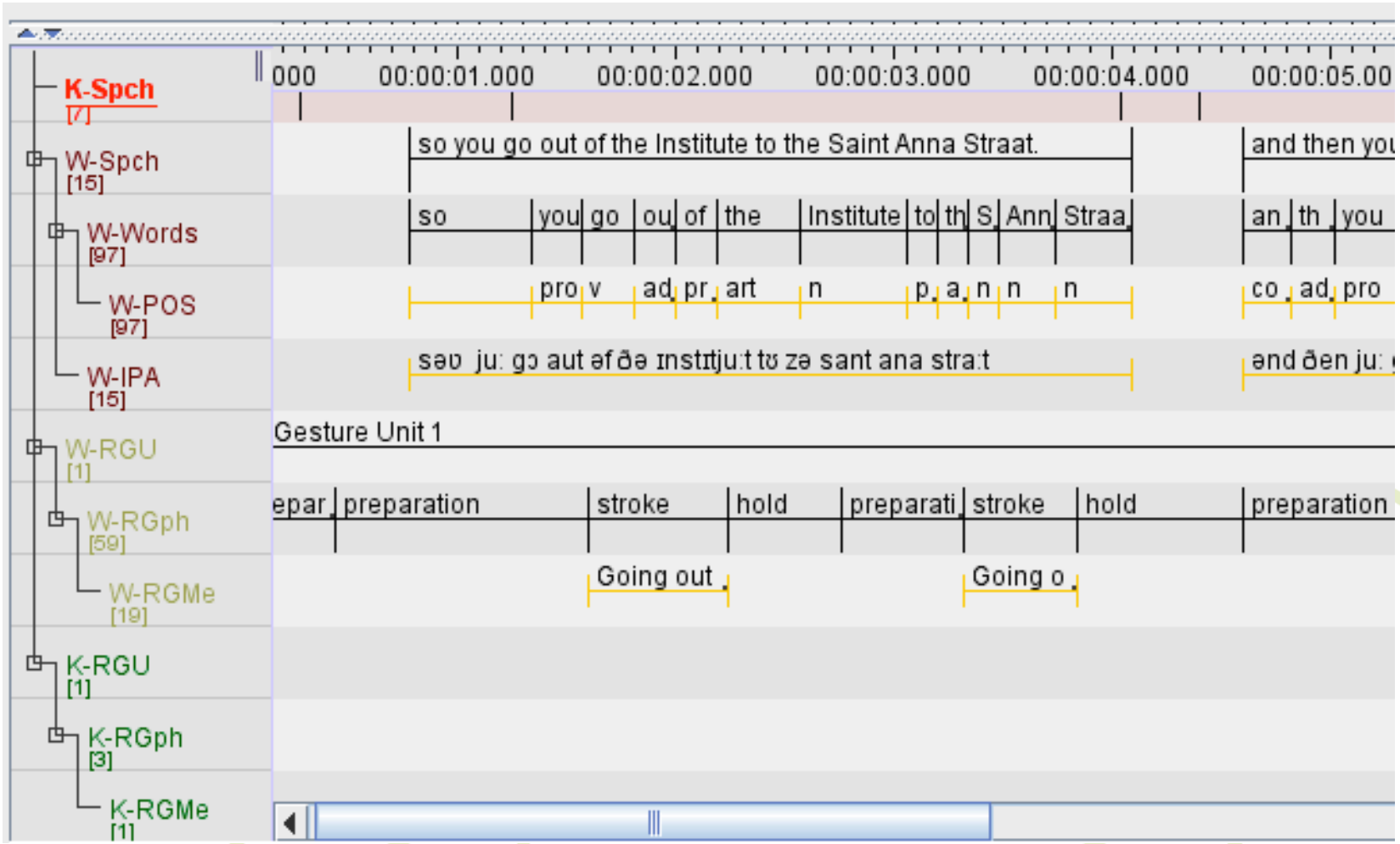
Tiered approach - 2



Tiered approach - 3



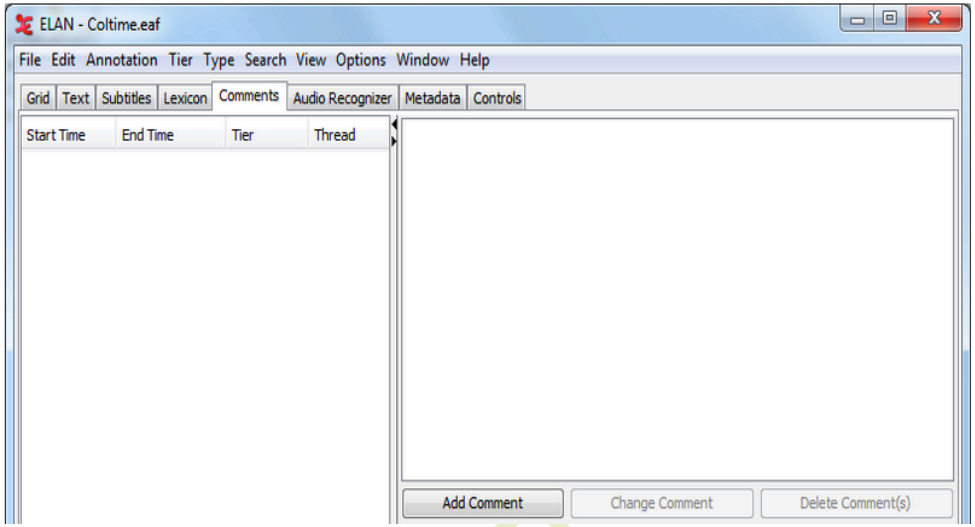
Tiered approach - 4



Annotating in ELAN - master the controls

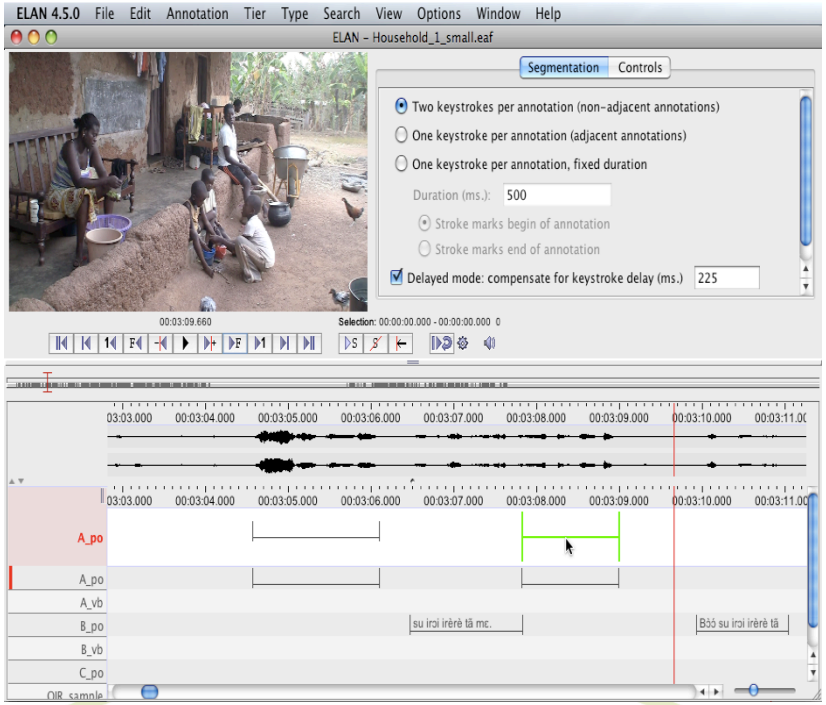
ELAN understands "annotation" as the content of the segments (*e.g., transcripts, codes, etc.*):

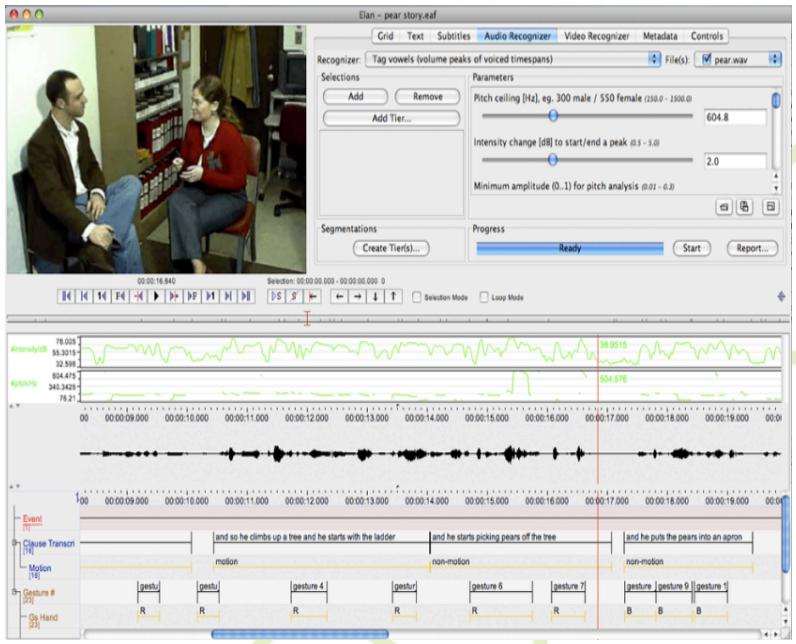
- **segments** are containers for annotations
- Each annotation is entered on a **tier** and assigned to a **time interval**
- There is also the possibility to add **comments**

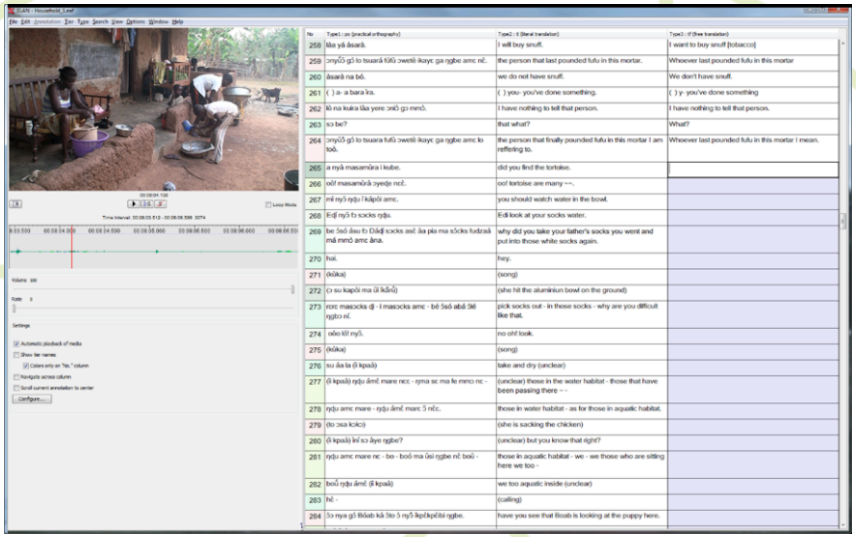


ELAN's working modes

- Segmentation mode
- Annotation mode
- Synchronization mode
- Transcription mode

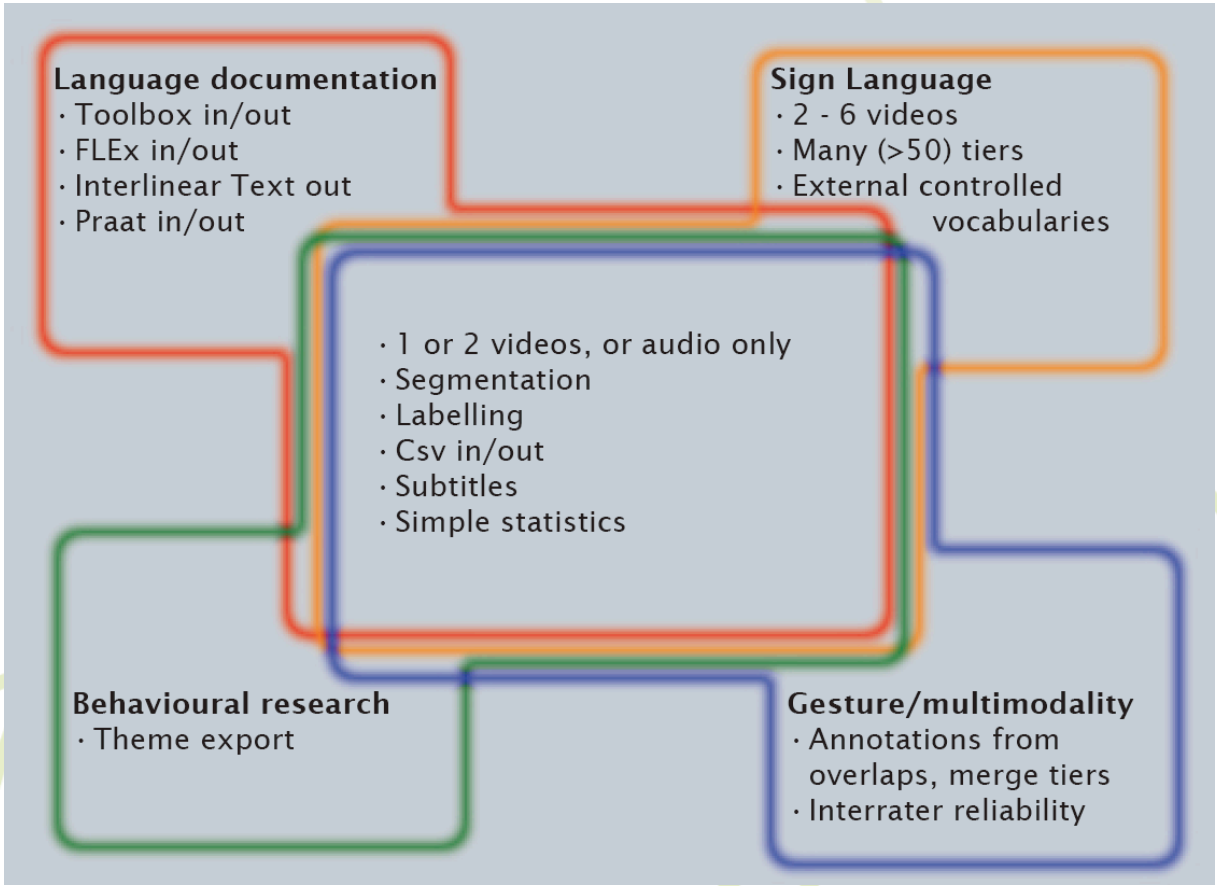






ELAN user communities

- Language documentation
- Sign Language research
- Gesture research
- Multimodality research
- Behavioural studies
- Psychology, psychiatry
- Other...



Original slide by Han Sloetjes

Linguistic Analysis - Free Tool

SketchEngine at https://auth.sketchengine.eu/#login

User Manual at
https://www.sketchengine.eu/user-guide/user-manual/

Landscape

What affective social sign processing scholars do with spoken data

Social Signal Processing

"How can emotions be generated in computers, **recognized** by computers, and expressed by computers?"

"SSP is the computing domain aimed at the modeling, **analysis**, and synthesis of social signals in human-human and human-machine interactions."

"How can we automatically **analyse** and **interpret** human behavior in human-human and human-machine interactions?"

Social Signal Processing in Oral History

Khiet Truong's Video:

https://drive.google.com/file/d/1-ooofaXzR3RWe9fL75NBQGT34B0KdbH2/view

Location:

https://drive.google.com/drive/folders/150Qe3itZArR0s2sAArPcaFUJT2YI9Y_9

Appreciating differences

- Scholars often ingrained in their ways
- But can see potential in alternative methods

- Can provide additional new angles e.g. for social scientists to explore the role of audio/ speech/emotion in a guided conversation

- Social science/oral historians - better question the role of an acquired textual 'transcription'

Challenge no 1: the jargon

Jargon - techniques and tools

- Getting to know a ‘new paradigm’
- Time is needed to work through analytical approaches and their own dedicated terminology
- So many acronyms, unfamiliar terms!
- Annotation types: annotation, coding, nodes
- Different and competing metadata schemas
- Suggest lay descriptions of tools and key functionality

Challenge no 2: accessing the technology

Technology - installation and (lack) of documentation

- Time and sometimes expertise may be needed to install open-source software, register, and get it working
- Technical barriers e.g. the platform and OS - linux, mac, windows only
- Many researchers struggle to download software, suggesting a lack of basic technical proficiency
- And the problem of Admin access!

Challenge no 3: using the tool

The tool - working the tool

- Time is needed to work through a new tool/ become familiar with it
- Preprocessing of data may be needed
- Watching an orientation video in advance helps
- Step-by-step guides and familiar teaching data and work through exercises very useful
- Be brave and curious
- Ask for advice and/or training
- Work **collaboratively** across discplines

Challenge no 4: finding and accessing data

The data - much interview data not open or discoverable

- Hard to find suitable accessible collections for the user
- Limited ‘metadata’, often manually created, overview. Free text and index terms **often describe the research/data collection method**, not content
- Data publishers could use **language tools to better emerge content** e.g. auto-summarisation and concept extraction

Thank you!

corti@essex.ac.uk
@LouiseCorti

ukdataservice.ac.uk
oralhistory.eu

Transcription Chain

Christoph Draxler
Arjan van Hessen
Henk van den Heuvel

PARTHENOS Workshop
Oct. 8, 2019
Sofia

Overview

Why?

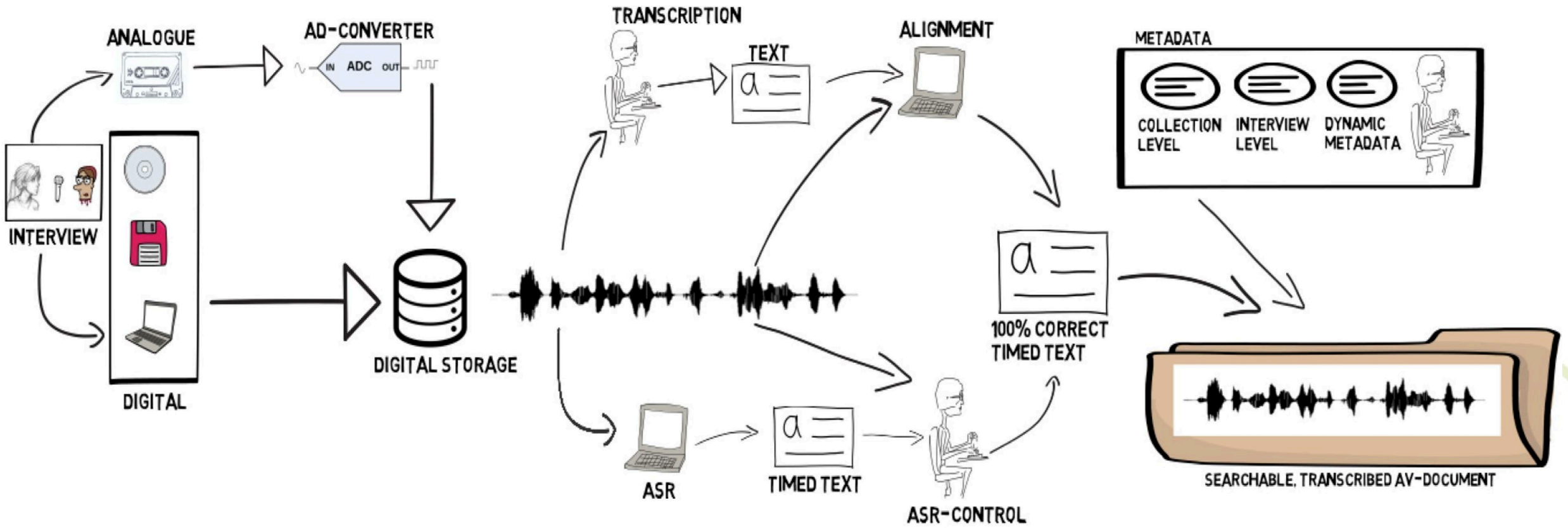
There’s no free lunch

It’s not rocket science

You get what you paid for

Life is beautiful

Transcription Chain



Why?

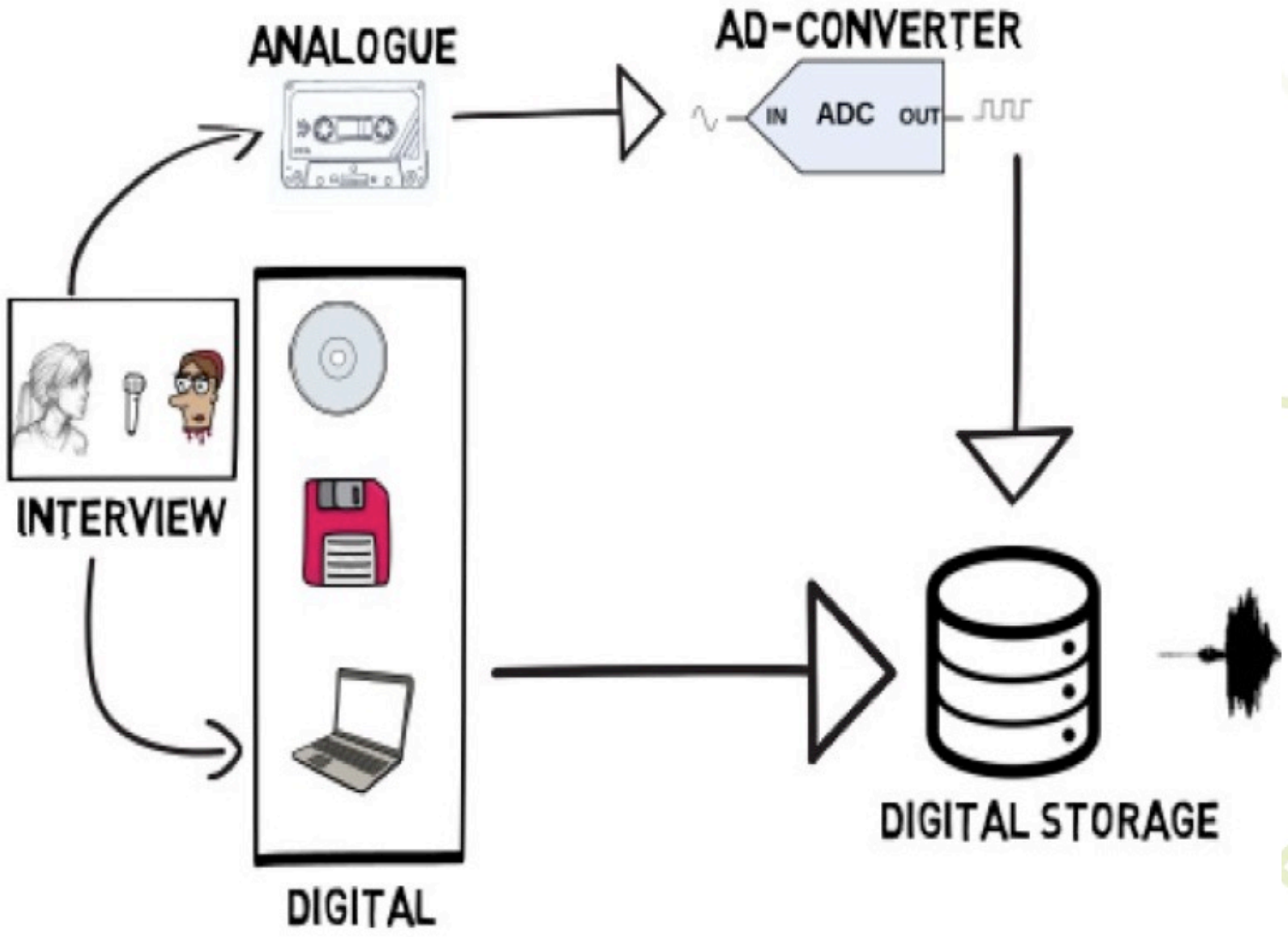
Transcripts are only one source of data

Audio (and video) signals carry additional information

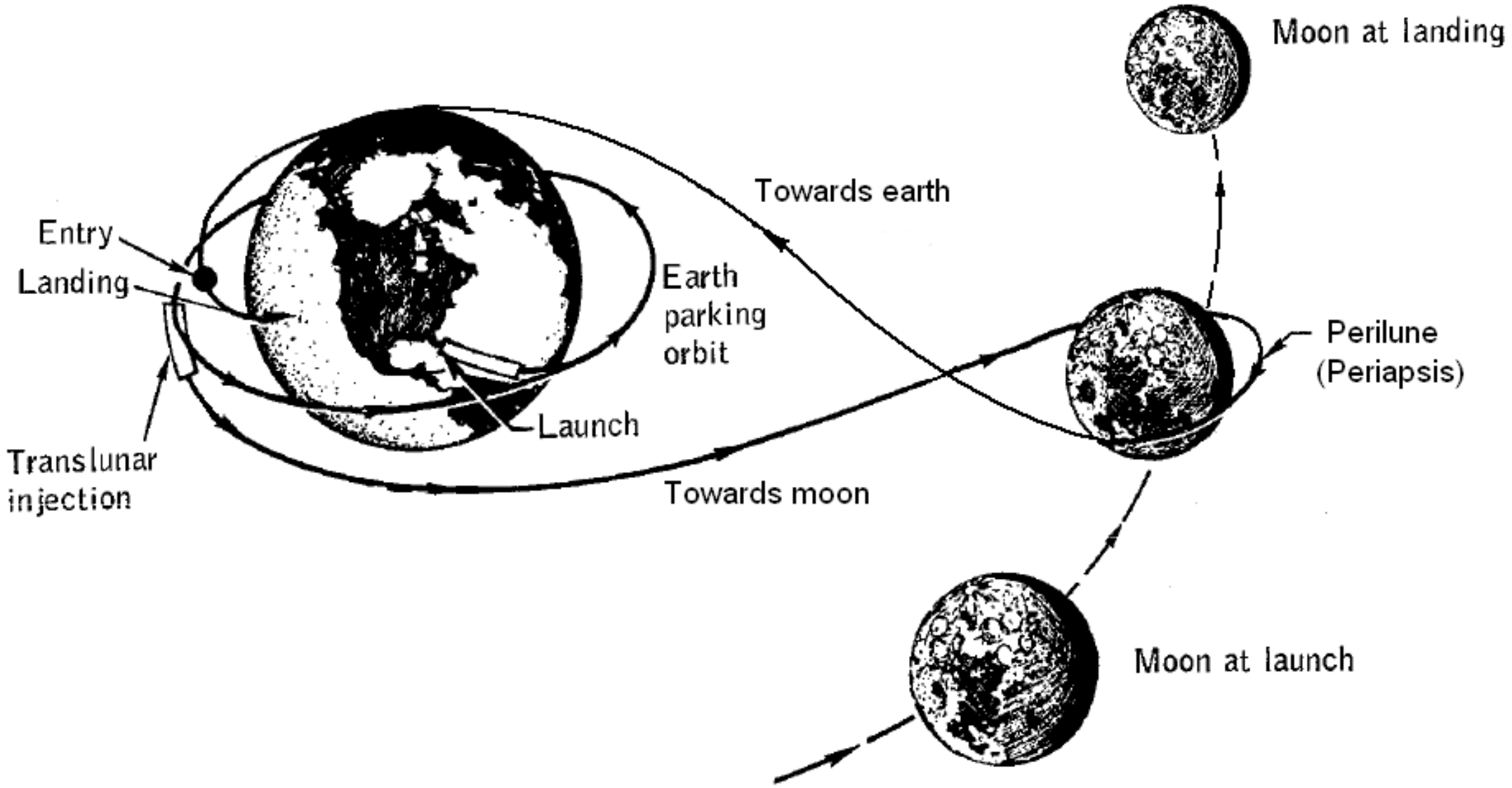
- Speaker sex, age, health
- Dialect, regional accent
- Emotional state
- Paralinguistic information
- Discourse analysis
- Recording environment

There‘s no free lunch

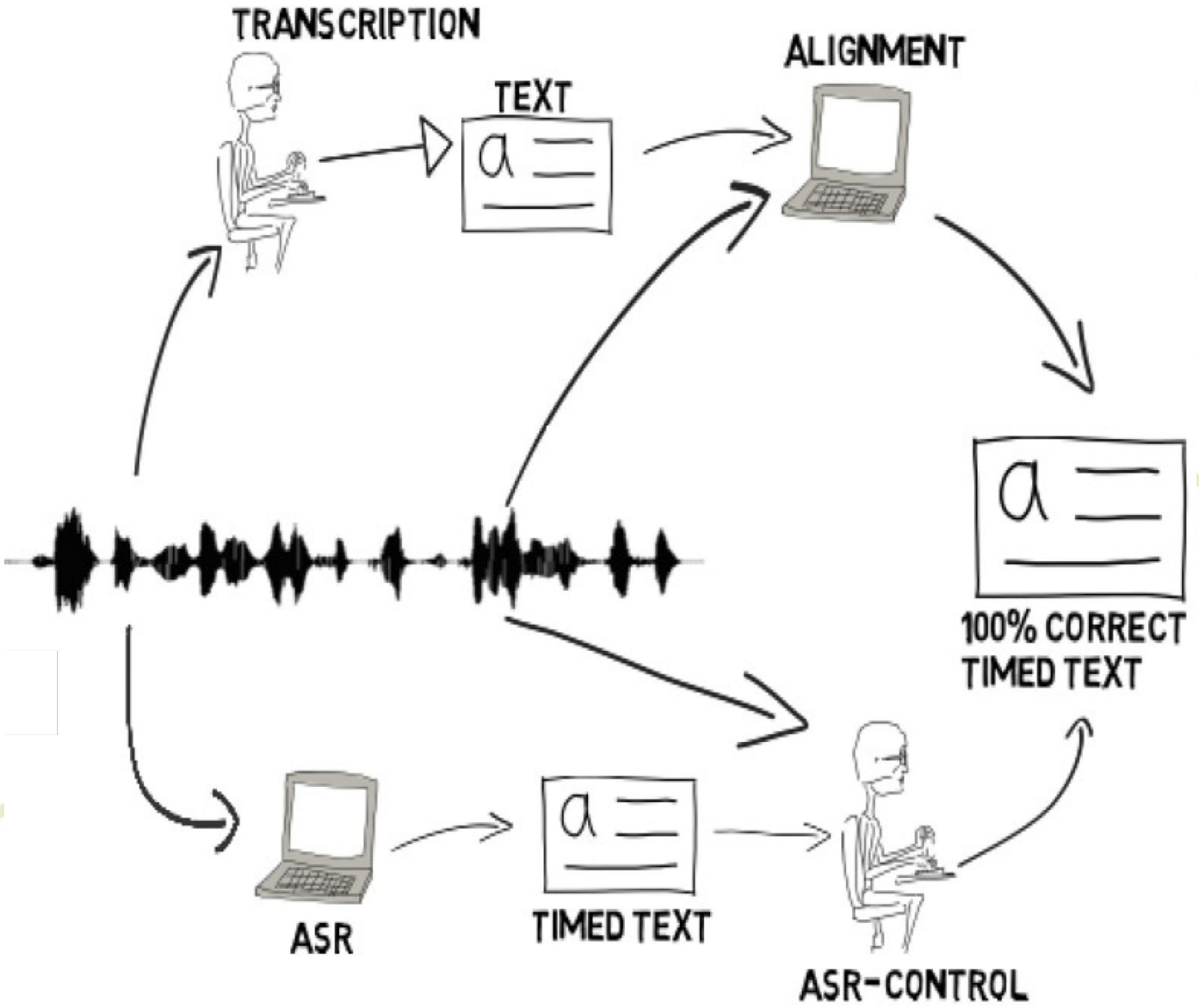
Data Collection and Preprocessing



It‘s not rocket science

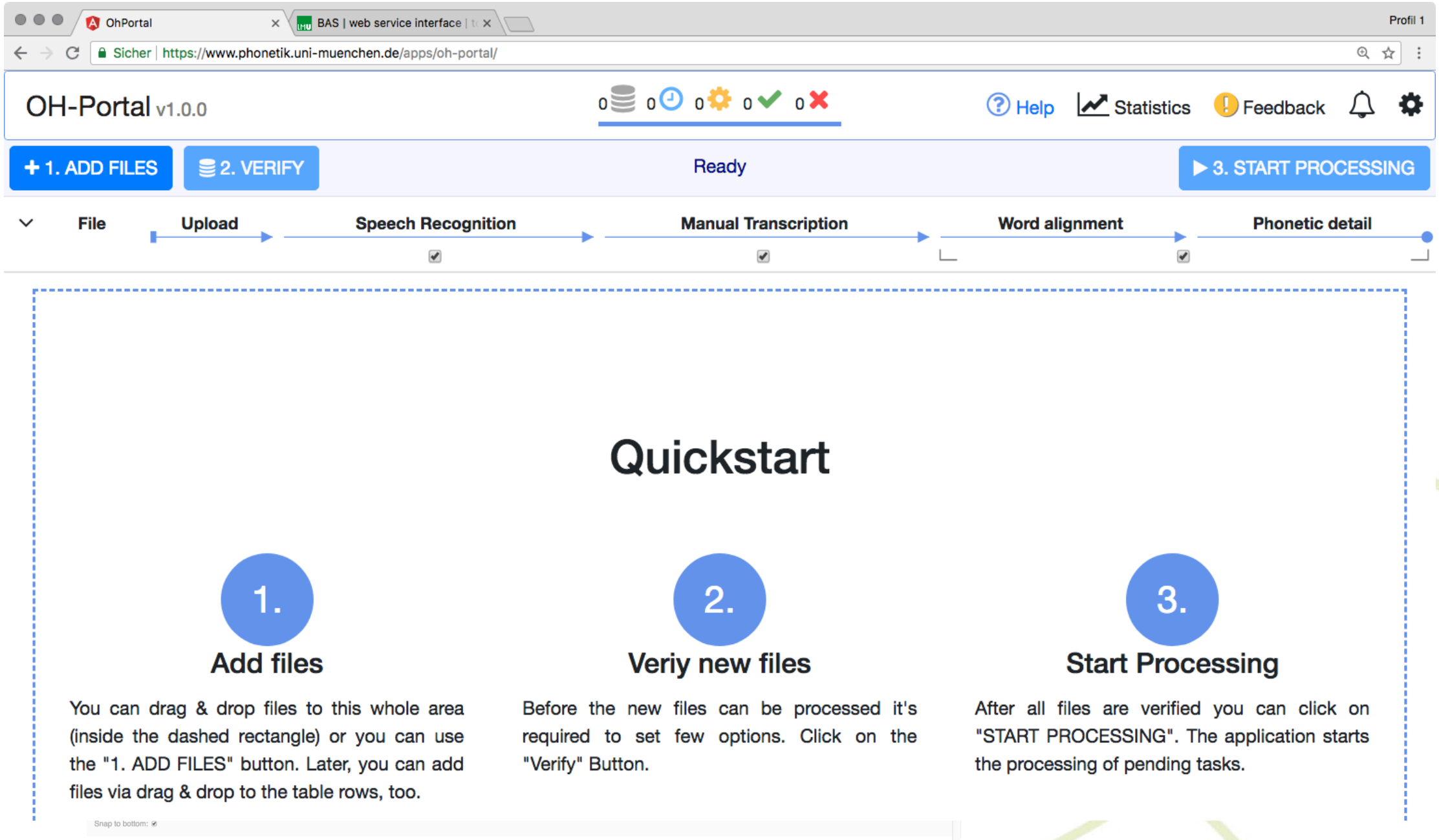

Moon at landing
Towards earth
Entry
Landing
Earth
parking
orbit
Perilune
(Periapsis)
Launch
Translunar
injection
Towards moon
Moon at launch

…it‘s much more difficult

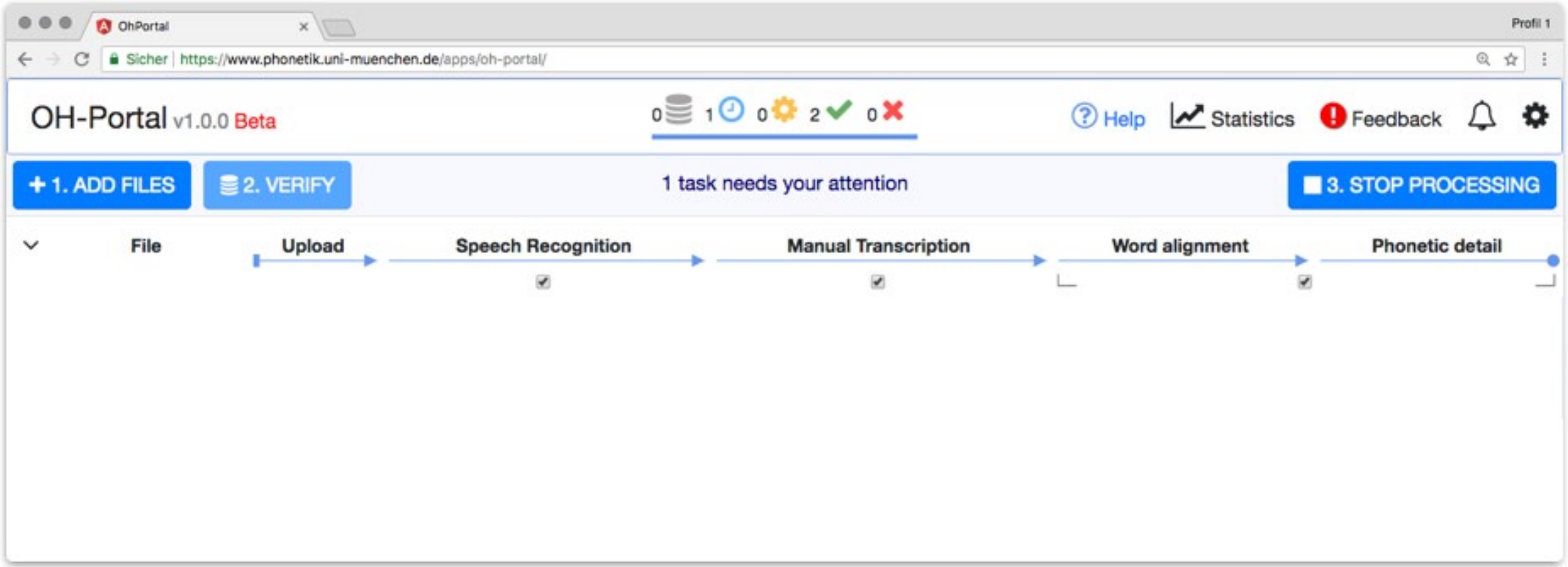


Processing Steps

- Verify and upload files
- Select language and ASR provider
- Manual correction of transcription
- Word alignment
- Data analysis

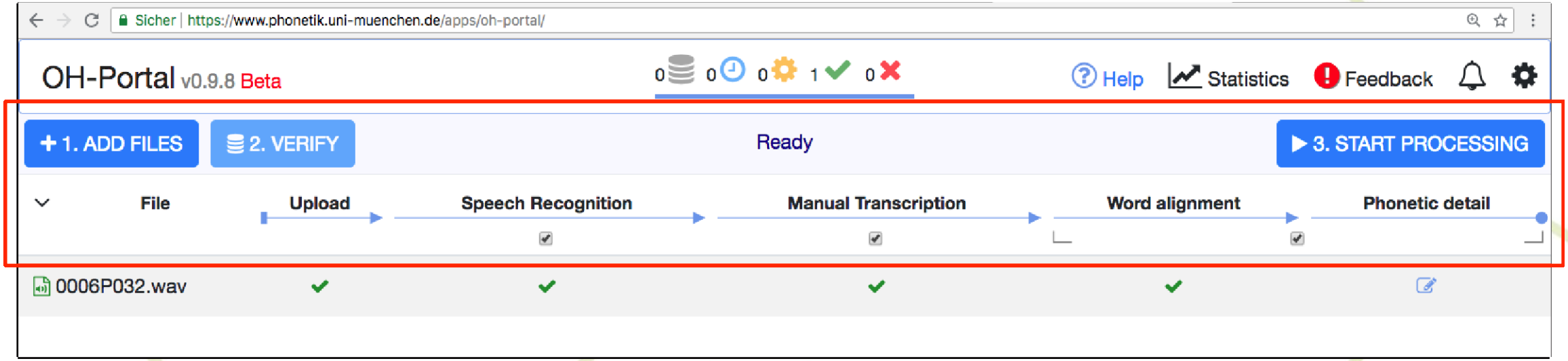

OhPortal
BAS | web service interface
Profil 1
Sicher | https://www.phonetik.uni-muenchen.de/apps/oh-portal/
OH-Portal v1.0.0
0 0 0 0 0
Help
Statistics
Feedback
+ 1. ADD FILES
2. VERIFY
Ready
▶ 3. START PROCESSING
File
Upload
Speech Recognition
Manual Transcription
Word alignment
Phonetic detail
Quickstart
1.
Add files
You can drag & drop files to this whole area (inside the dashed rectangle) or you can use the "1. ADD FILES" button. Later, you can add files via drag & drop to the table rows, too.
2.
Veriy new files
Before the new files can be processed it's required to set few options. Click on the "Verify" Button.
3.
Start Processing
After all files are verified you can click on "START PROCESSING". The application starts the processing of pending tasks.
Snap to bottom:

OH portal page

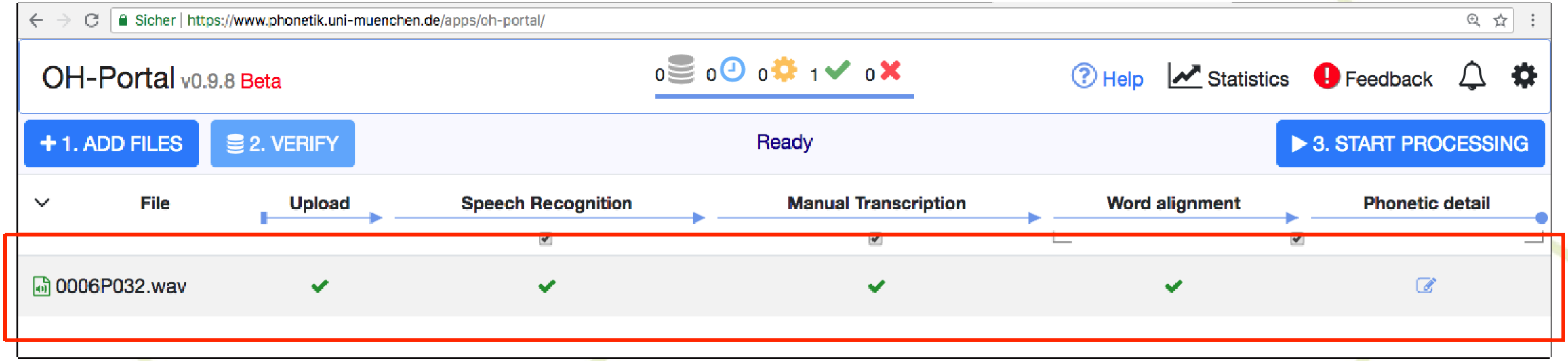


https://www.phonetik.uni-muenchen.de/apps/oh-portal/

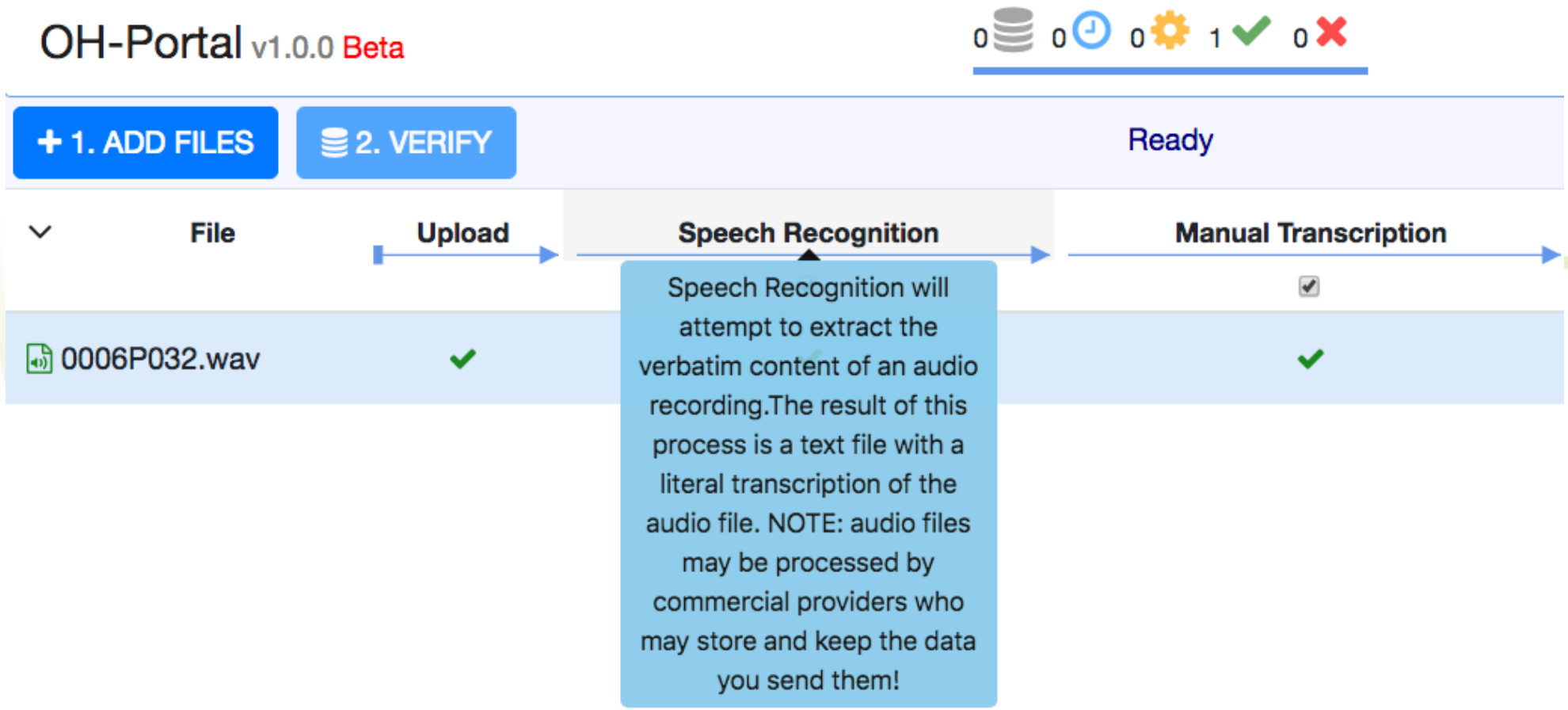
Processes & control



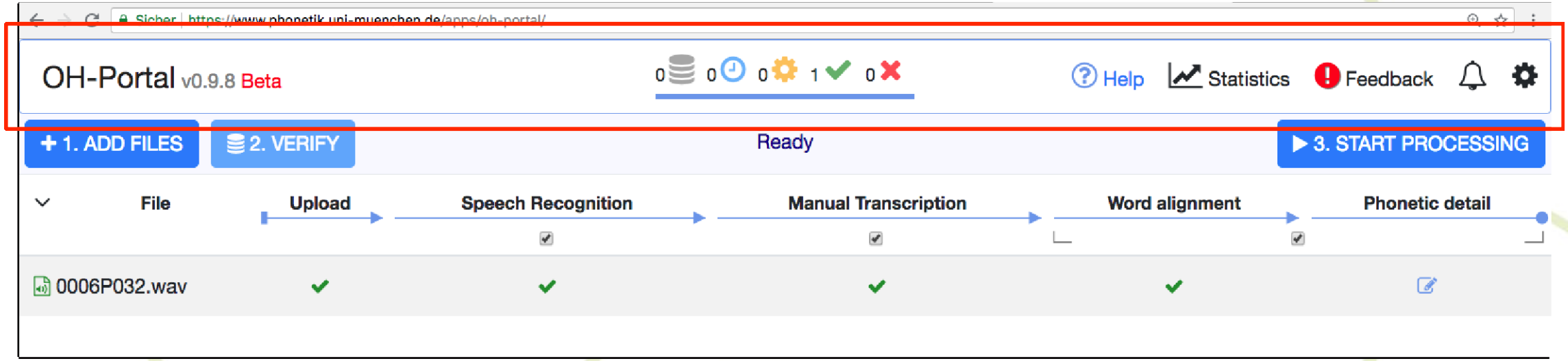
Status table



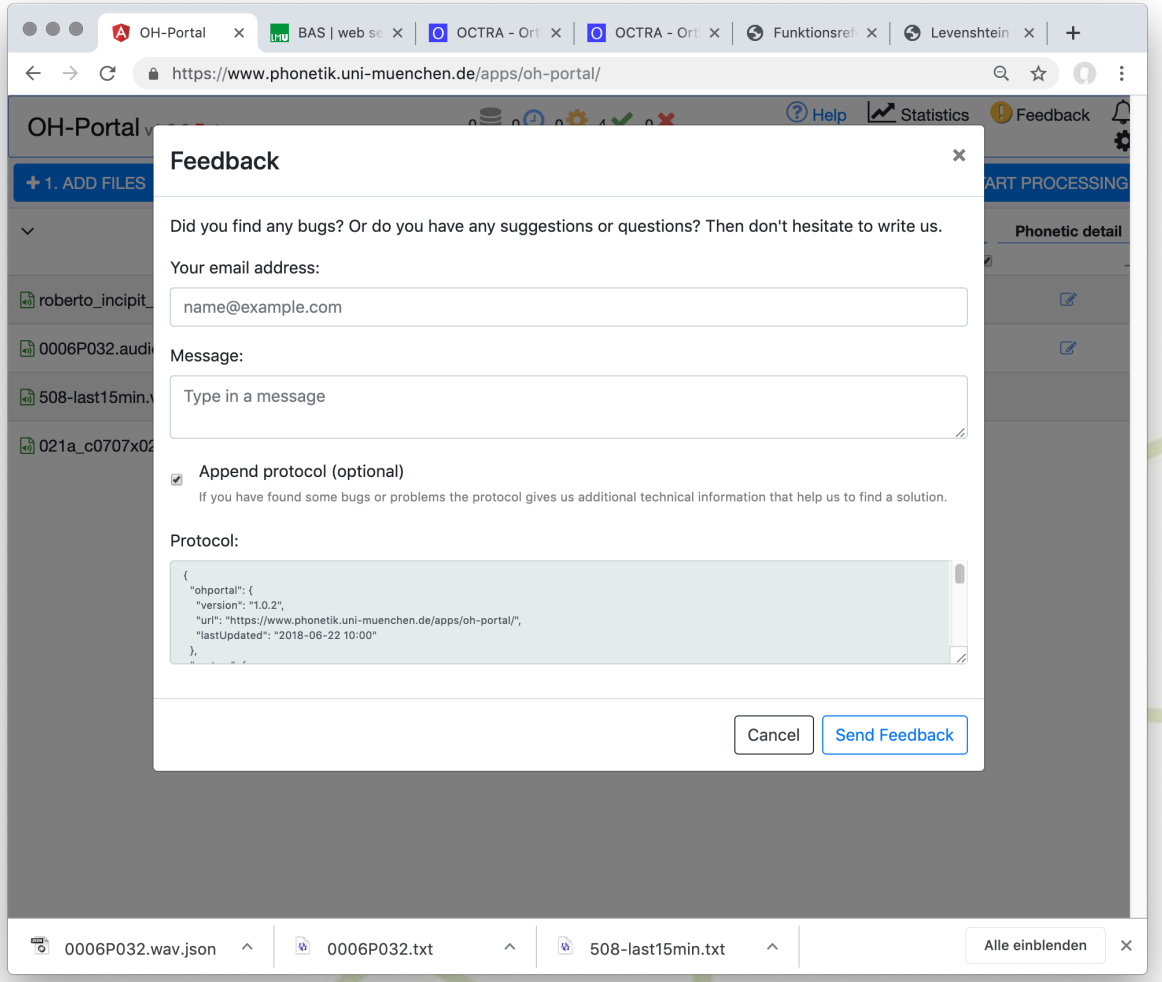
Tooltips


OH-Portal v1.0.0 Beta
0 0 0 1 0
+ 1. ADD FILES
2. VERIFY
Ready
File
Upload
Speech Recognition
Manual Transcription
0006P032.wav
Speech Recognition will attempt to extract the verbatim content of an audio recording.The result of this process is a text file with a literal transcription of the audio file. NOTE: audio files may be processed by commercial providers who may store and keep the data you send them!

Feedback & information



Feedback



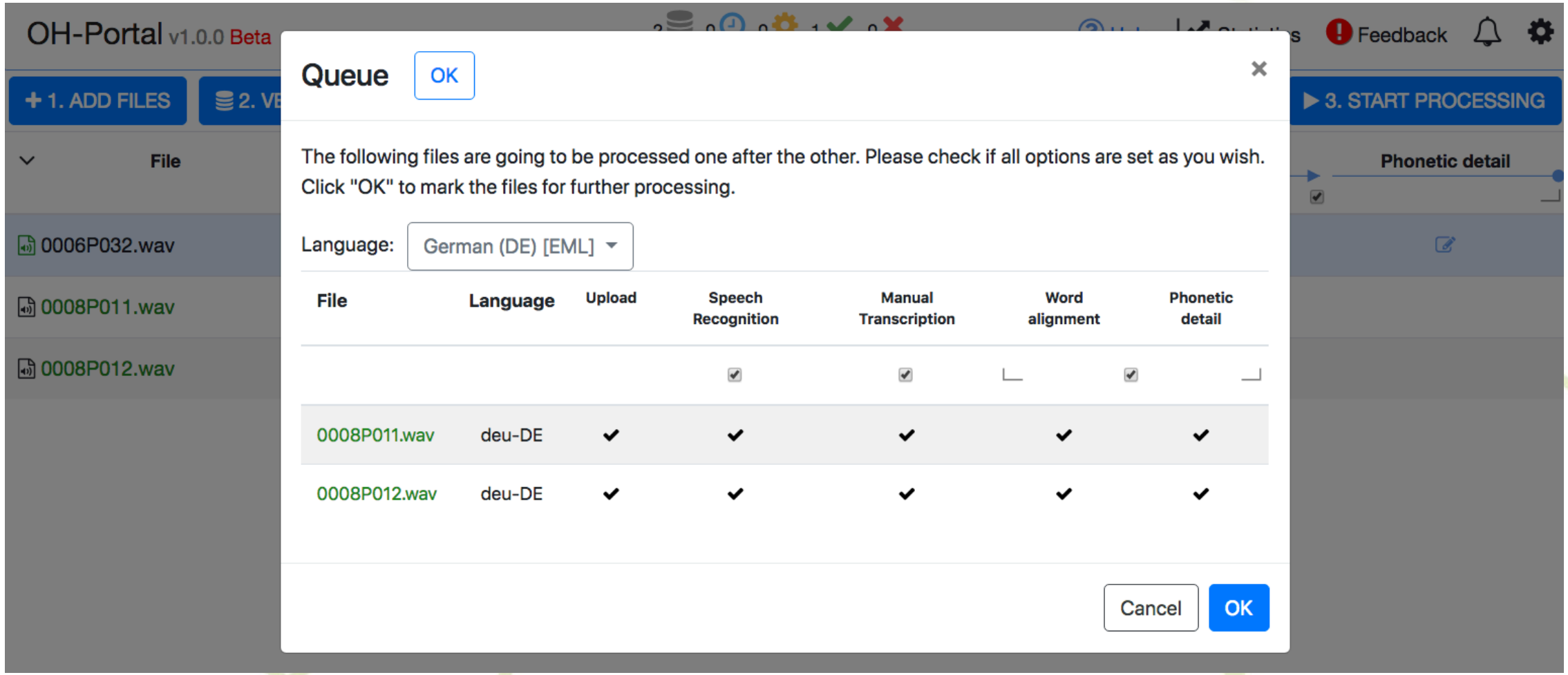
Your feedback is important

- Compiles vital system information
- Text input for problem description
- Maps to our server logs to facilitate tracing errors

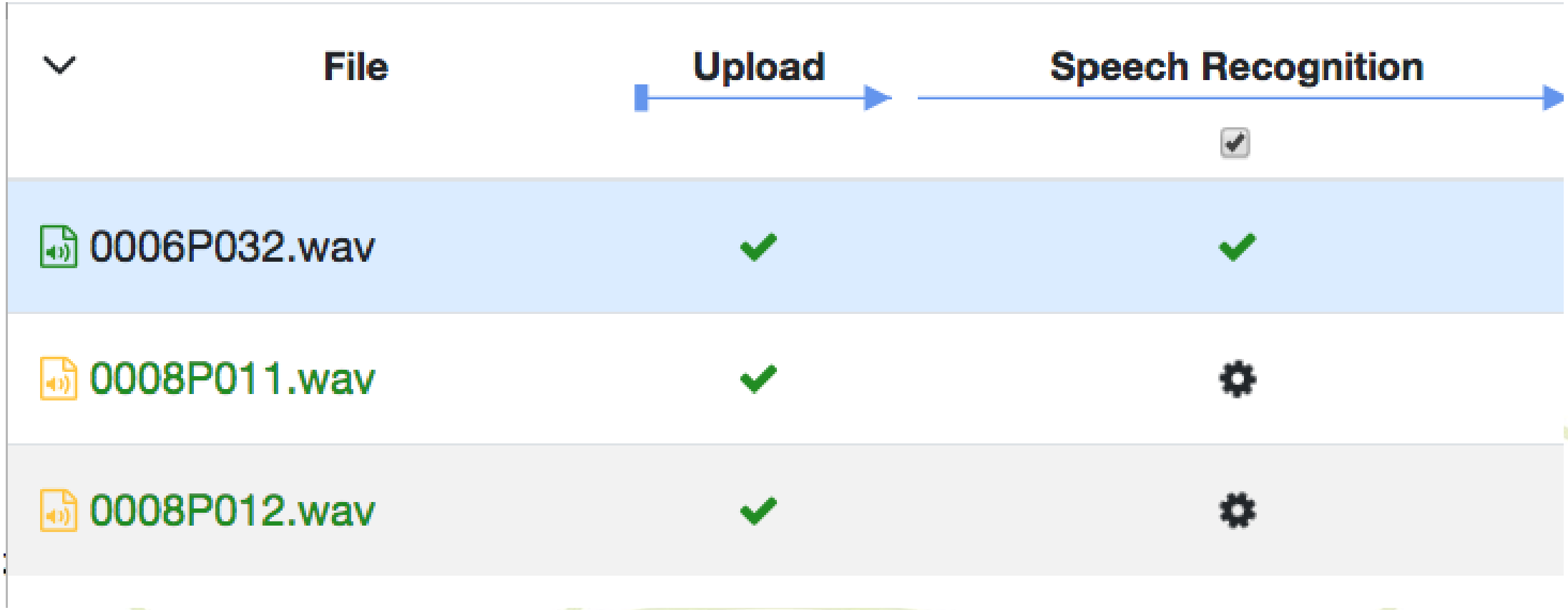
Verify audio files

- Drag & drop audio files to browser window
- Click on „Verify“ to perform basic tests
- Optional steps have a check box (default: active)
- Note: some restrictions apply
  - all recordings must be in the same language
  - only WAV-files (mono or stereo) are supported
  - tested on Chrome and Firefox browsers
  - currently only **English, German, Dutch, Italian**

Verify files settings & upload

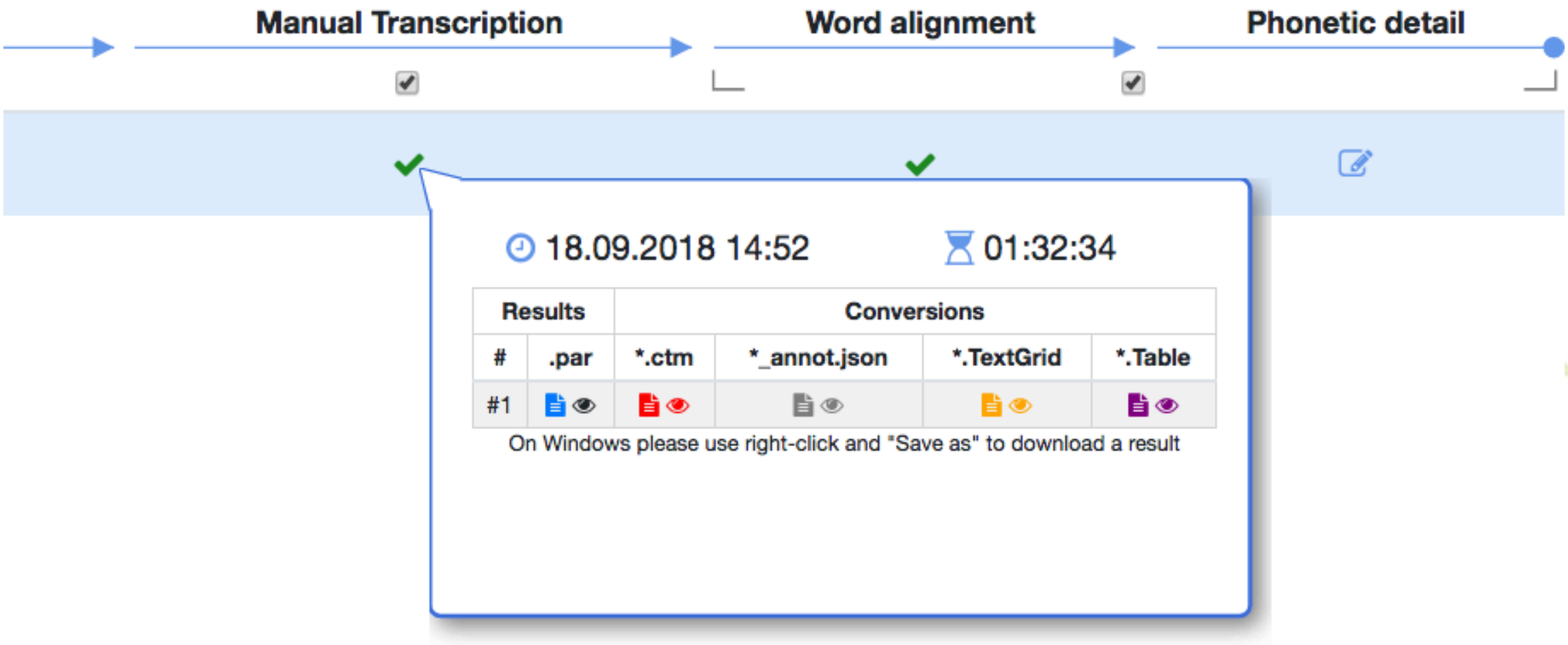


Automatic speech recognition (ASR)



error success

View & download intermediate files



You get what you paid for

Levenshtein distance

- Standard measure for the similarity of strings
- Based on edit operations
- counts the number of
  - insertions
  - deletions
  - substitutions
- to convert string A to string B

In the following slides, the Levenshtein distance is used to measure the difference between ASR outcomes

- **Example:** Levenshtein distance between the character sequences

  my argument
  and
  mike yeah argument

  is **7**
- substitute ‚y‘ by ‚i‘
  insert ‚k‘ and ‚e‘
  insert ‚y‘, ‚e‘, ‚a‘, ‚h‘

Why manual transcription?

Manual transcript

my uhm argument for useful history is that actually all history starts by being oral because somebody goes and tells somebody about something and then they write it down yeah okay now back to so so the guy writing it down making it wrong or the guy may tell it wrong so just because it is written down it doesn't mean it is true uh I mean I I have commanded a brigade in a war the Falklands war I know how for example war diaries are kept where every unit stop me telling you something you know already to produce an official diary of what they were doing every day okay and that is logged now in the national archive yeah now some academic historian looking at diaries being like the new tablet of stone brought down by by Moses it must be true because it's written down I believe across war diaries where there are actual lies being written because often a war diary would be written to uhm place the writer and others in the best possible light yeah ok so is not necessarily accurate at all secondly things like uh after action reports a pilot coming back from a mission debriefing the intelligence officer so this is what I did okay it might be backed up by photographs and things like that but it is still an oral account patrols coming back will report to to the intelligence officer an oral account so without wishing to sort of bore you I actually think oral history has a huge value provided you know what you are doing when you are listening to it and you have a very good idea of who the people are and therefore how trustworthy or other they are I think before I bore you to tears and and

Manual transcript is considered the **gold standard**

Why manual transcription? (English results)

Google

my argument for useful history is actually all history start by being horrible because somebody doesn't tell someone about something and they write it down ok now so so the guy writing it down make it wrong with the guy make him wrong so just because it's written it doesn't mean it's true I have commanded a Brigade in a war the Falklands War I know how data is written that the British waters where is unit stocking something every unit has produced an official diary of what they were doing everyday ok and that's launched in the archive academic and look up on war Diaries being like them tablets of stone brought down by Moses it must be true that's not I've come across war Diaries where they are that relies been written was often awarded it'll be written to place the writers and authors in the best possible light ok so it's not necessarily accurate at all secondary things like after action report a pilot earn from the pilot is actually…

IBM Watson

Mike yeah argument the use of all history is actually all history stop by the wall because somebody doesn't tell someone about something they write down yeah okay now in writing it down maybe it won't go it alone just because it's written down it doesn't mean it's true uh I mean I I cannot be again in the war the fog of war and I know how for example water is correct a quarter of a unit yeah well it can in fact has produced an official direct what they were doing every day in the law now in the in in the national park yeah no some of the local border if being the the brought down by but in truth written that not in the water is where they are actually like being written the water in the region to I'll place the writers and all in the best possible light yeah Kay that's not very accurate at all secondly things like half dressed report of miss in the end of the day …

Why manual transcription? (English results)

Google

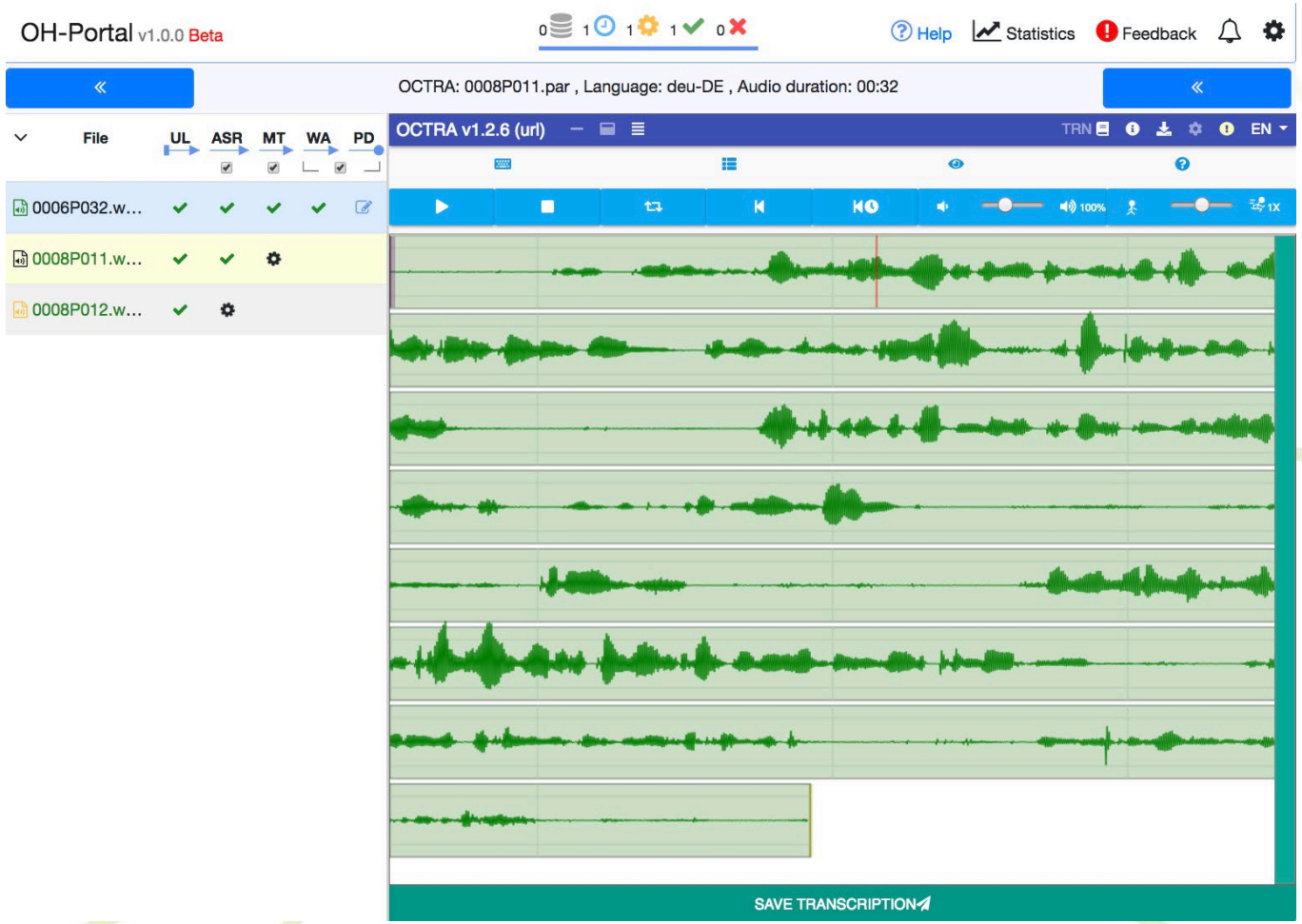
my argument for useful history is actually all history start by being horrible because somebody doesn't tell someone about something and they write it down ok now so so the guy writing it down make it wro… it's … com… kno… whe… pro… eve… aca… tabl… true… the… it'll… the… acc… report a pilot earn from the pilot is actually…

IBM Watson

Mike yeah argument the use of all history is actually all history stop by the wall because somebody doesn't tell someone about something they write down yeah okay now in writing it down maybe it …n in …e …n in …e …n …e …he …half

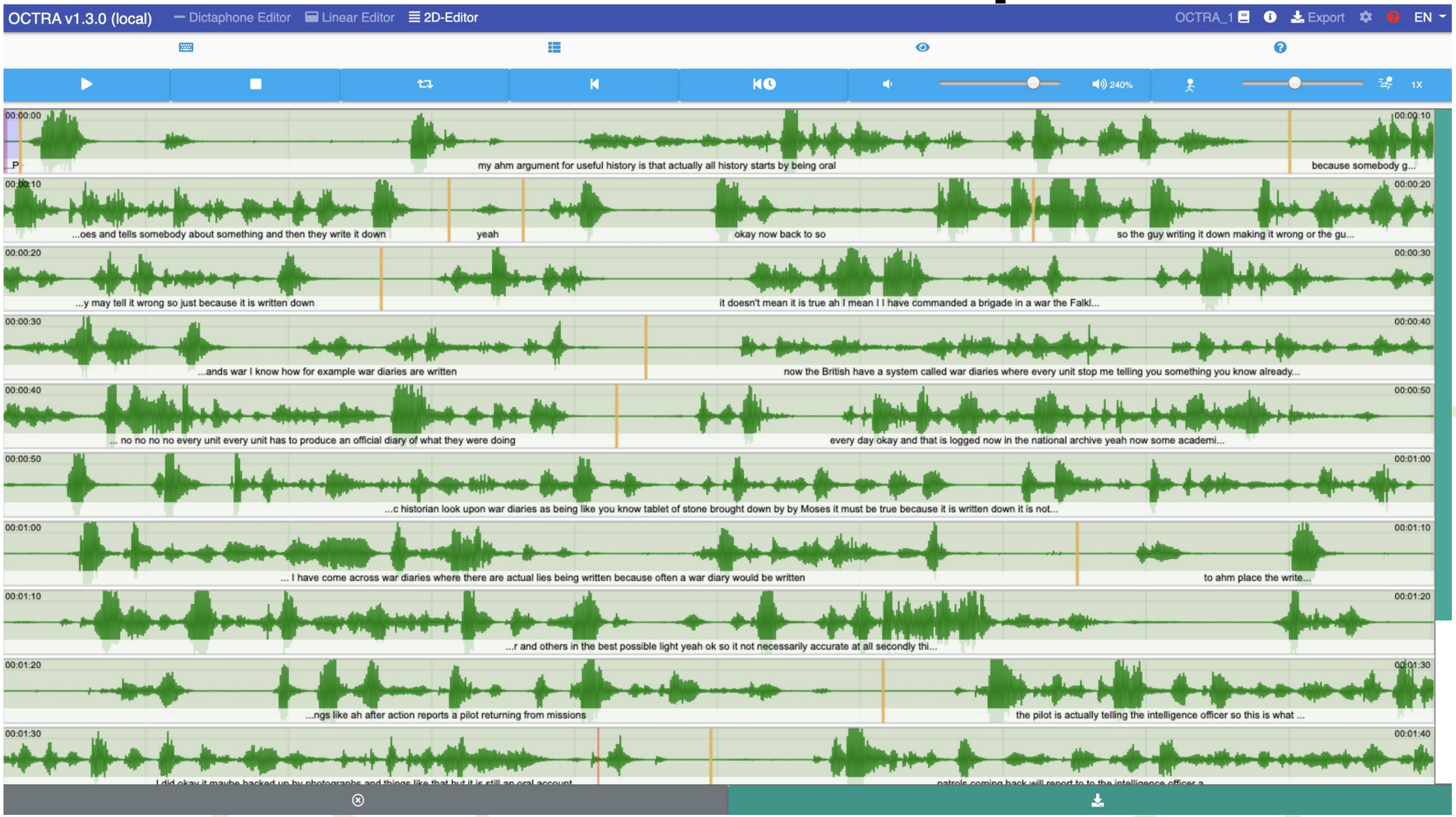
Type	manual	Google	IBM
Characters	1723	1395	1173
Words	332	256	238
Levenshtein distance (chars)		481	802
error rate		27.9%	46.5%
Levenshtein distance (words)		150	213
error rate		45.2%	71.2%

Manual transcription

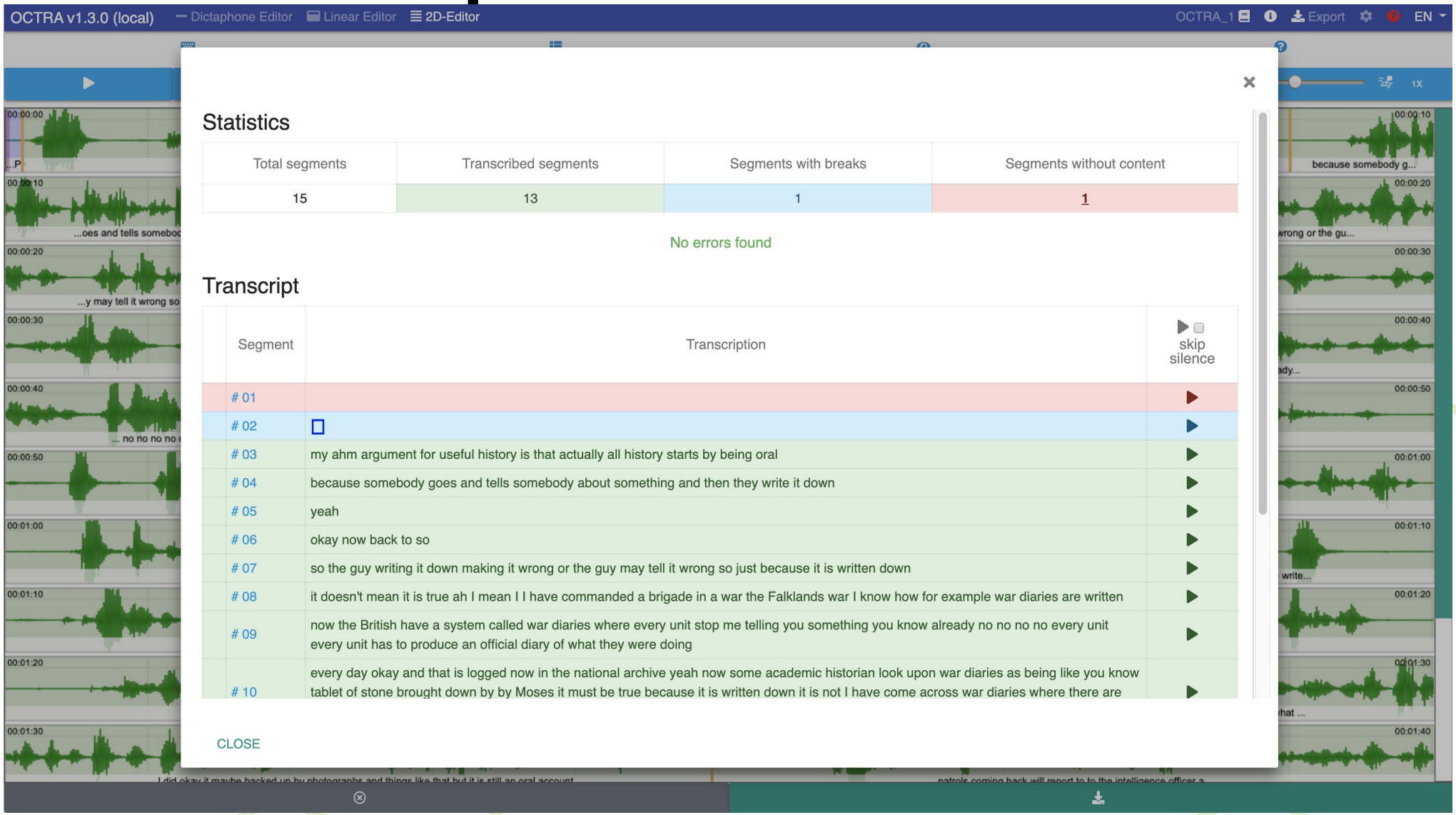


Check (and correct) the output of the ASR service

Listen to and correct transcript



Save transcript



ASR vs. manual transcription pilot study

- 10 audio recordings
- monologues on ‚communication‘
- 3-5 min long
- video camcorder in seminar room
- 2 transcribers
  - 5 ASR + manual correction,
  - 5 fully manual transcripts
- measure

$$tfactor = \frac{dur_{transcription}}{dur_{recording}}$$

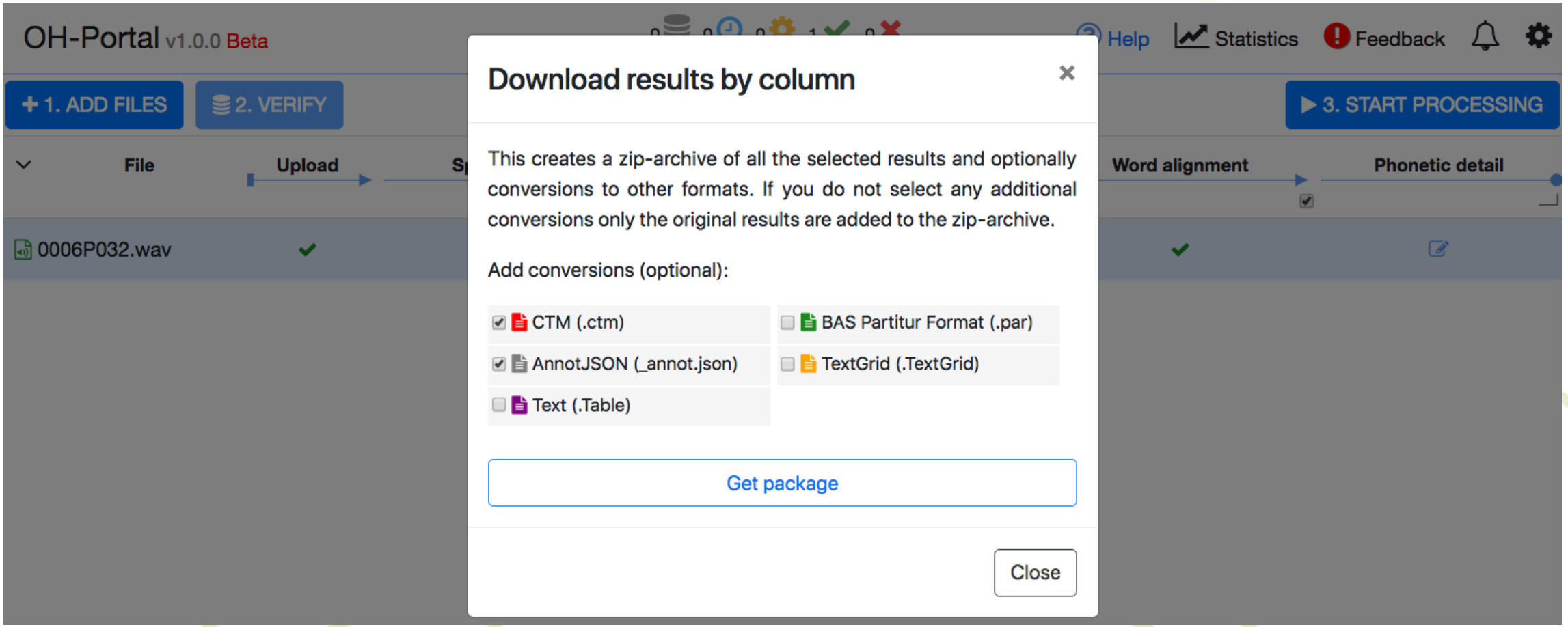
Results II: ASR vs. manual transcription

type	tfactor
ASR + manual correction	8.52
fully manual transcription	9.43

Expect higher tfactors for real-world data (interviews, historical recordings, dialectal speech, etc.)

Life is beautiful

Download results column-wise



Where to go from here?

Format	Tool	Application
CTM	ASR	speech recognition transcripts
EAF	ELAN	multi-level audio & video annotation
AnnotJSON	Emu SpeechDB, any programming language	multi-level phonetic and linguistic analysis, statistics, programming
BAS Partitur	WebMAUS	multi-level (word, phoneme) forced alignment
TextGrid	Praat	phonetic analysis, signal processing
plain text	any text editor	content or pattern search
table	Excel, R, DBMS	content statistics