

SSHOC Webinar

CLARIN Hands-on Tutorial on Transcribing Interview Data



SPEAKERS



Christoph Draxler

Researcher, Institute of Phonetics and
Speech Processing at Ludwig Maximilian
University Munich

draxler@phonetik.uni-muenchen.de



Henk van den Heuvel

Director Centre for Language and Speech
Technology, Radboud University
Nijmegen, Faculty of Arts

H.vandenHeuvel@let.ru.nl



NOTES

-  **The webinar is being recorded.** All participants will receive a link to the recording later today.
-  **Slides are available:** See the chat box for the link.
-  **Questions?** Put them in the chat box. We'll put questions to the speakers at the end of the webinar.

PREREQUISITES

Eduroam account

 Please make sure your eduroam account works – you will need it to access the OH Portal

Audio sample files

 OH_Portal_demo.wav test run of the webinar

 interview_Ben_Tucker.wav interview with Ben Tucker

 Thompson-01.wav interview with British general

 historic_interview.wav analog socio-historical recording

Google Chrome browser



Project:



SSHOC

social sciences & humanities open cloud



Horizon 2020
European Union Funding
for Research & Innovation

Type of action & funding:
Research and Innovation action
(INFRAEOSC-04-2018)

Partners: 45

(20 beneficiaries + 25 LTPs)

SSH ESFRI Landmarks and Projects
& international SSH data infrastructures

Project budget:
€ 14,455,594.08

Duration: 40 months
(January 2019 – 30 April 2022)

Project website:
www.SSHopencloud.eu



Objectives:

- creating the social sciences and humanities (**SSH**) part of European Open Science Cloud (**EOSC**)
- maximising **re-use** through **Open Science** and **FAIR** principles (standards, common catalogue, access control, semantic techniques, training)
- interconnecting existing and new infrastructures (clustered cloud infrastructure)
- establishing appropriate **governance model** for SSH-EOSC

EXPECTED IMPACT



The Social Sciences and Humanities are seamlessly integrated in the European Open Science Cloud



Availability of an EU-wide, easy-to-use SSH Open Marketplace, where tools and data are openly accessible



EU-wide availability of high quality "cloud ready" SSH tools and high quality SSH data



EU-wide availability of trusted and secure access mechanisms for SSH data, conforming to EU legal requirements



State of the art Research Infrastructure in several pilot domains advanced through dedicated SSH data pilots cluster projects



Maximising reuse through Open Science and FAIR principles (standards, common catalogue, access control, semantic techniques, training)

CLARIN in six bullets

- **CLARIN** is the Common Language Resources and Technology Infrastructure
- **ESFRI** ERIC status since 2012, Landmark since 2016
- that provides easy and sustainable access for scholars in the **humanities and social sciences** and beyond
- to **digital language data** (in written, spoken, video or multimodal form)
- and **advanced tools** to discover, explore, exploit, annotate, analyse or combine them, wherever they are located
- through a **single sign-on** online environment

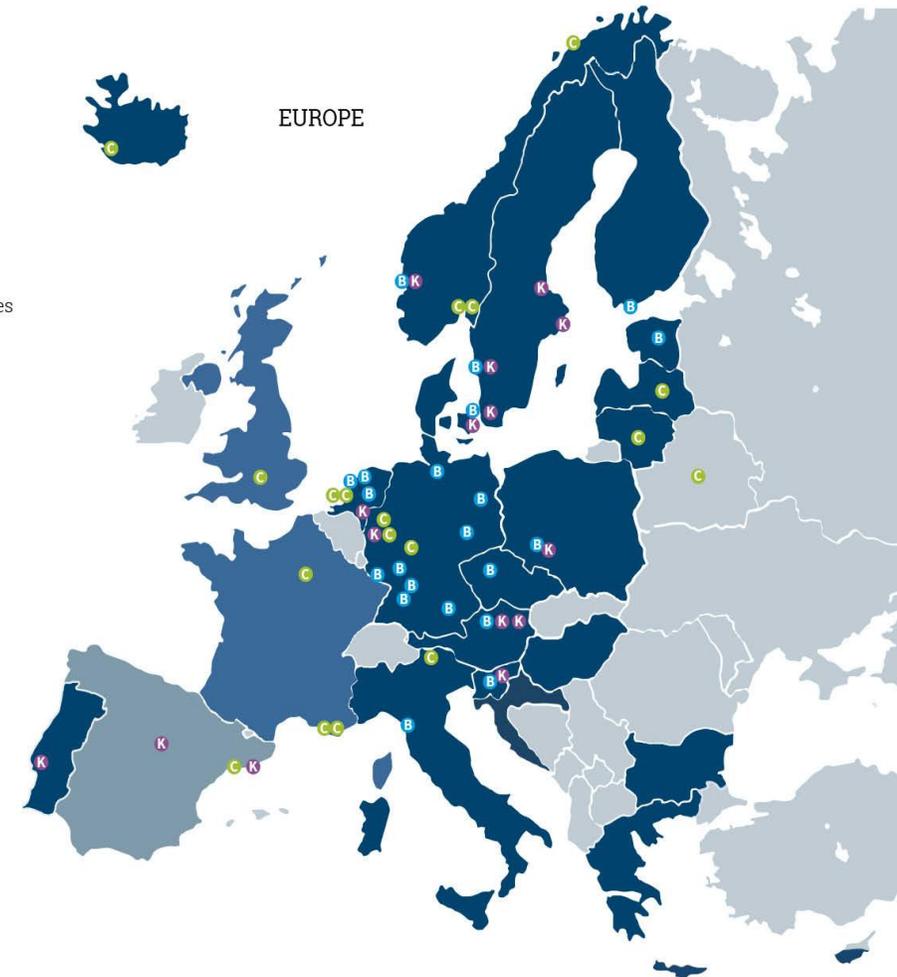
CLARIN ERIC in members and centres

A consortium of:

- 21 members: AT, BG, CY, CZ, DE, DK, EE, FI, GR, HR, HU, IS, IT, LT, LV, NL, NO, PL, PT, SE, SI
- 3 observers: FR, UK, ZA
- >50 centres



- ERIC members
- Observers
- Countries with participating centres
- B Centre Providing Data
- C Centre Providing Metadata
- K Knowledge Centre



Technical pillars

- **Federated identity** - letting users login to protected data and services with their own institutional login and password
- **Persistent identifiers** - enabling sustainable citations of electronic resources
- **Sustainable repositories** - digital archives where language resources can be stored, accessed and shared
- **Flexible metadata and **concept definitions**** - to ensure semantic interoperability when describing language resources
- **Content search** - offering a search engine for a wide range of language resources
- **Web service chaining** - giving users the possibility to freely combine language processing services

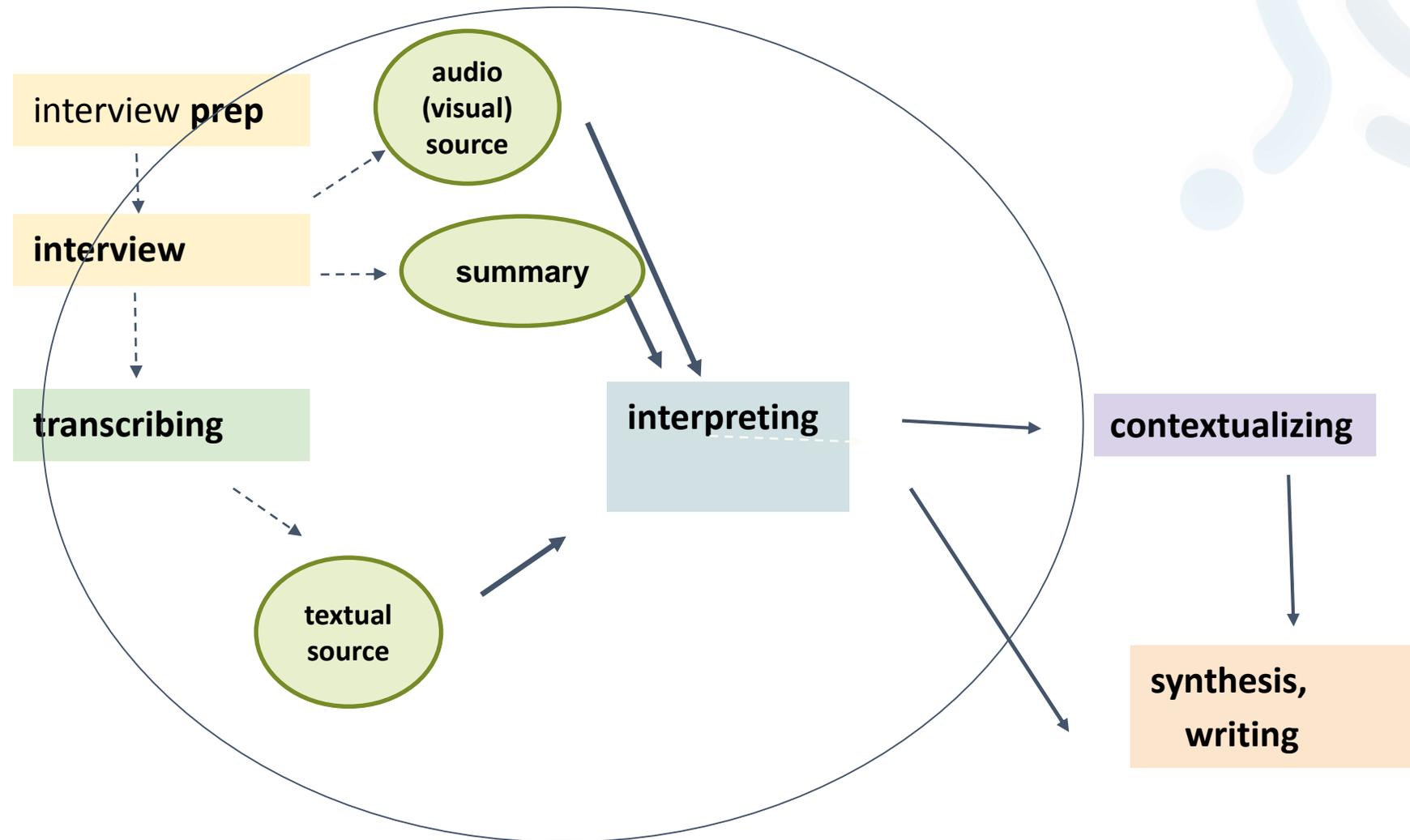
CLARIN & ORAL HISTORY



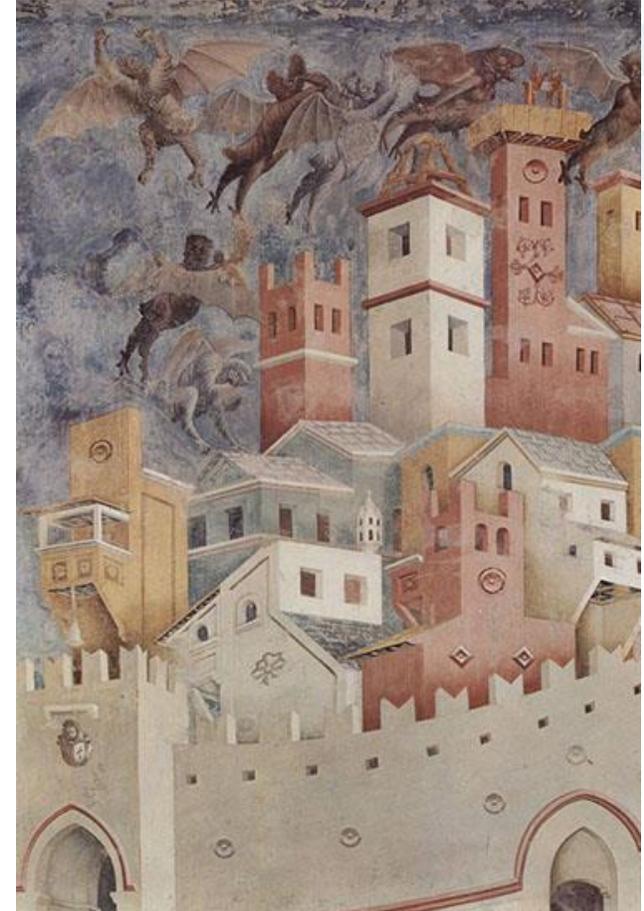
- 🌀 From Wikipedia, the free encyclopedia:
 - 🌀 Oral history is the collection and study of historical information about individuals, families, important events, or everyday life using audiotapes, videotapes, or transcriptions of planned interviews. These interviews are conducted with people who participated in or observed past events and whose memories and perceptions of these are to be preserved as an aural record for future generations.
- 🌀 Central are interview data



What do oral historians do with interviews?



INTERVIEWS & LANGUAGE AND SPEECH TECHNOLOGY



🌀 Task force with CLARIN support:
<https://oralhistory.eu/>

CLARIN workshop in Arezzo in May 2017

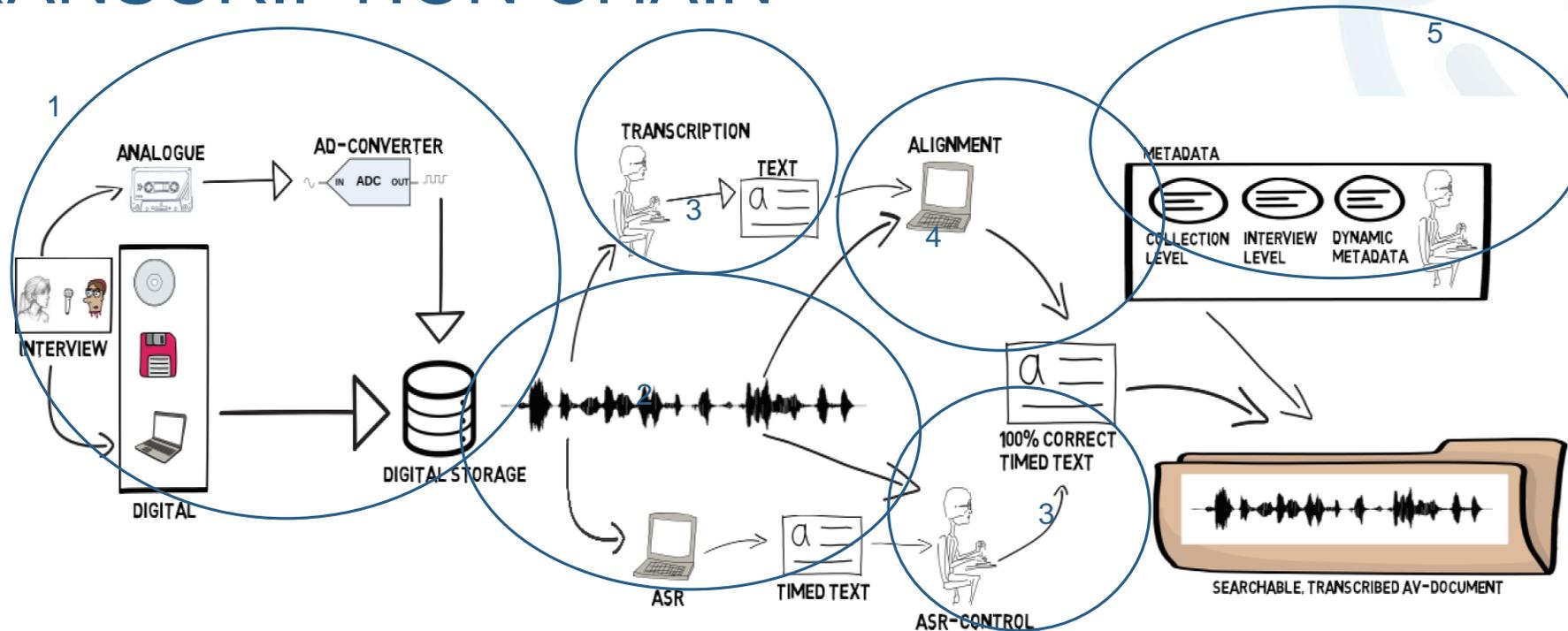
Goals:

- 🌀 Draft a proposal for establishing an infrastructure directed at the orthographic transcription of existing OH-collections using Automatic Speech Recognition (ASR)
- 🌀 Show and practice various building blocks
 - 🌀 **Before and during workshop**
- 🌀 Proposal for integration of building blocks into CLARIN infrastructure

Direct Result

- 🌀 Setting up the Transcription Portal at BAS Munich
- 🌀 Basis for a follow-up CLARIN workshops including (qualitative) analyses for which transcriptions are a precondition

THE TRANSCRIPTION CHAIN



1. Analog-Digital Conversion (ADC)
see <https://oralhistory.eu/showcases/transcription-chain#digitizing-analogue-recordings>
2. Automatic Speech Recognition (ASR)
3. Transcription correction
4. Synchronisation of Audio & transcription (Alignment)
5. Metadata



THE TARGET GROUPS

Starting in Arezzo, 2017:

- 🌀 Historians and social science users who conduct research with recorded interview data sources;
- 🌀 Linguists who use spoken language sources;
- 🌀 Software tools specialists who develop support data processing and analysis tools.

Participants in next workshops (Munich, 2018; Utrecht, 2019) further included:

- 🌀 Computational linguists
- 🌀 Sociolinguists
- 🌀 Social sign processing scholars
- 🌀 Psychology, including psycholinguistics and mental health studies



THE TARGET GROUPS

But think also of:

- 🌀 Law scholars
- 🌀 Social economists
- 🌀 Empirical religion scholars
- 🌀 Pharma-medical scientists

- 🌀 Etc. Etc: everyone using interviews in their work

- 🌀 What is your background?



Transcription Chain and the OH-Portal

- 🌀 Transcription types
- 🌀 Introduction to the T-Chain
- 🌀 Demo: OH-Portal
- 🌀 Recommended workflow



Transcription Types

- 🌀 Raw transcript
- 🌀 Raw Transcript with diarization
- 🌀 Transcript with some markup
- 🌀 Broad transcript
- 🌀 Man vs. Machine I & II



Raw text transcript (Thompson-01)

my uhm argument for useful history is that actually all history starts by being told and then they write it down yeah okay the guy writing it down making it wrong or the guy may tell it wrong so just doesn't mean it is true uh I mean I I have commanded a brigade in a war for example war diaries are written now the British have a system called war diaries that prevents me from telling you something you know already no no no no every unit has an official diary of what they were doing every day okay and that is lodged in a file now some academic historian look upon war diaries as being like you know the tablets or stone brought down by by Moses it must be true because it is written down it is not I have come across war diaries where there are actual lies being written because often a war diary would be written to uhm place the writers and all these others in the best possible light yeah ok so it not necessarily accurate at all secondly things like uhm after action reports uh pilots returning from missions the pilot is actually telling the intelligence officer this is what I did okay it maybe 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 and it's an oral account so without wishing to sort of bore you I actually think that 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

This type of transcript is often used to evaluate the ASR recognition rate, or as input to other tools.



Raw Transcript diarization

S1 my uhm argument for useful history is that actually all history starts by someone about something and then they write it down

S2 yeah

S1 okay now

S2 back to to see this is

S1 so so the guy writing it down making it wrong or the guy may tell it wrong mean it is true uh I mean I I have commanded a brigade in a war the Falklands are written now the British have a system called war diaries where every unit know already

S2 no no no no

S1 every unit every unit has to produce an official diary of what they were in the national archive

S2 yeah

S1 now some academic historian look upon war diaries as being like you know the tablets of stone brought down by Moses it must be true because it is written down it is not I have come across war diaries where there are actual lies being written because often a war diary would be written to uhm place the writers and all these others in the best possible light

S2 yeah

S1 okay so it not necessarily accurate at all secondly things like uhm after action reports uh pilots returning from missions the pilot is actually telling the intelligence officer this is what I did okay it maybe 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 and it's an oral account so without wishing to sort of bore you I actually think that 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

This type of transcript is often used

- **to perform automatic processing of the individual speakers' contribution,**
- **for content analysis, and**
- **as the basis for in-depth analyses**



Transcript with some markup

S1 my <F/> argument for useful history is that actually all history starts by being someone about something and then they write it down

S2 yeah

S1 okay now

S2 back to to see this is

S1 so so the guy writing it down making it wrong or the guy may tell it wrong so j mean it is true <F/> I mean I I have commanded a brigade in a war the <NEGEO example war diaries are written now the British have a system called war diaries from telling you something you know already</META>

S2 no no no no

S1 every unit every unit has to produce an official diary of what they were doing every day okay and that is <VOCAB>lodged</VOCAB> in now in the national archive

S2 yeah

S1 now some academic historian look upon war diaries as being like you know the tablets of stone brought down by by <NEBIB>Moses</NEBIB> it must be true because it is written down it is not I have come across war diaries where there are actual lies being written because often a war diary would be written <P 0.97/> to <P 0.46/> <F/> <P 0.96/> place the writers and all these others in the best possible light

S2 yeah

S1 okay so it not necessarily accurate at all secondly things like <F/> after action reports <F/> pilots returning from missions the pilot is actually telling the intelligence officer this is what I did okay it maybe 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 and it's an oral account so <META>without wishing to sort of bore you</META> I actually think that 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 <META>I think before I bore you to tears and and</META>

This type of transcript is often used to mark regions of interest for further research, e. g. discourse analysis, emotion analysis



Broad transcript

My argument for useful history is that actually all history is written down because somebody goes and tells someone about something and they write it down.

Just because it is written down it doesn't mean it is true. For example, war diaries. The British have a system called war diaries where they produce an official diary of what they were doing every day. Historians look upon war diaries as being true because they are written down. Not! I have come across war diaries where there are errors because often a war diary would be written to place the writers in the best possible light. Secondly, things like after action reports. Pilots or patrols returning from missions tell the intelligence officer what they did, maybe backed up by photographs, but it is still an oral account.

I actually think that 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 they are.

This transcript is

- **close to the original recording, and**
- **it features punctuation, text smoothing etc. to make it easy to read.**



Summary of Transcription Types



🌀 The transcription types differ in terms of

🌀 verbatim fidelity

🌀 target audience

🌀 reusability

🌀 human expertise vs. machine processing needed

🌀 In general, converting a

🌀 broad transcript to a more detailed one is not possible without the signal

🌀 very detailed transcript to a less detailed one may take as long as starting from scratch (unless tags can be removed automatically)

🌀 Thus, a raw transcript with diarization is the most flexible solution

Man vs. machine I (Thompson-01.wav)

my uhm argument for useful history is that actually all history starts by being oral because somebody goes and tells someone about something and then they write it down yeah okay now back to to see this is 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 written now the British have a system called war diaries where every unit do stop me from telling you something you know already no no no no every unit every unit has to produce an official diary of what they were doing every day okay and that is lodged in now in the national archive yeah now some academic historian look upon war diaries as being like you know the tablets of stone brought down by by Moses it must be true because it is written down it is not I have come across war diaries where there are actual lies being written because often a war diary would be written to uhm place the writers and all these others in the best possible light yeah ok so it not necessarily accurate at all secondly things like uhm after action reports uh pilots returning from missions the pilot is actually telling the intelligence officer this is what I did okay it maybe 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 and it's an oral account so without wishing to sort of bore you I actually think that 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

my argument for the use of all history is a history Starts by being oral somebody goes until someone about something down in a war I know how for a God of War Diaries are written that the water unit stop every unit has produced in the diary what they were doing everyday ok and that lodged in our archive some tablets of stone brought down by by Moses it must be true I've come across Ward areas where they are our lives being written award will be written to place the writers in the best possible light after action report returning from missions this is what I did buy photographs and things and a cunt coming back no value provided you know what you're doing when you're listening and you have a very good idea who the people are and trustworthy to tell something

Manual: 1779 chars
Google ASR: 717 chars
Levenshtein distance: 1075
~ 60.4 % error rate

Manual transcript vs. Google ASR (default model, GB-English, called from BAS web services 17.02.2020)

Man vs. machine II (OH_portal_demo.wav)

you now see some preliminary remarks they concern the application of the OH portal it's an application that runs in a browser window so as long as you don't close the browser you will always have access to the data that you've been working on which is quite nice if something goes wrong you know the data is still there ok now we need to add some files so I will go to the operating system and select some files here is one by an English general and you can see now you can see the page of the OH portal there are four areas there is the top area where you can see a number of symbols they show what's going on at the moment they give you access to the help system they show you some statistics nice graphs showing how much time was spent doing ASR and how much time was spent doing manual transcription if you find anything strange in the software if anything doesn't work just click on the feedback button and send us a message the little bell tells you that something that you can switch on desktop notifications some processes may take a while so if you don't want to sit and stare at your screen but do something similar and do something more interesting than just click on the bell and it will the system will automatically notify if something's changed finally you have this button for settings then there's the second row with three blue buttons add files verify and start processing the middle area is reserved for messages so there you can see what's happening or what's going on at the moment in this case it says one audio file waiting to be verified then you see the column headings the column headers file upload speech recognition manual transcription word alignment per default speech recognition manual transcription and word alignment and phonetic detail are checked this means that our processing chain will go all the way from uploading the file through speech recognition you will do some manual transcription or correction and then finally there will be a word alignment and even phonetic details you might have seemed there is a little notification that came up and speech recognition actually has finished now the system or the file has arrived at the processing stage manual transcription and the symbol doesn't change to a little cog wheel instead has this little edit symbol let's go one step back first if you click on the icon on the on the green check mark next to the speech recognition or in the speech recognition column you can see that there are a number of options available already you can have the results of the speech recognition in one of the formats presented here so par is the BAS the Bavarian Archive for Speech Signals Partiturfile CTM is a format that's quite common for machine processing for speech recognition annot-json is new format that allows you to link the different elements on different annotation tiers which is very useful for later on processing TextGrid is a format might know when you've been working with Praat it shows the data in different tiers and on every tier all the items are marked with the precise time information table is a format that is very well suitable for R or Excel or any kind of statistics package and finally text is the raw text if we click on this we can see what Google made out of the audio file ok so this way you go through your file it might take a while until you've corrected everything in this case Google speech recognition results are not too good for a number of reasons probably the audio quality isn't too good although we can and understand it very well perhaps the topic that the speech recognition and engine is not familiar with we don't know that's one of the problems we access external services we have no idea what they actually offer and they only provide a very limited set of configuration options with which we could adapt the services to our needs ok I will close this and you can see the whole transcript let's imagine that the transcript was ok and we're done with the transcripts so all we do now is press on Send transcription meaning that now we have finished the manual transcription it is being sent to our server to the BAS server and now word alignment starts again this takes a few minutes and if you really have a very true manual transcript of what was said word alignment should give you very good results for the actual alignment of the words it will give you more it will give you the alignment of the individual sounds but maybe this is something you don't need for your kind of research but I think the word alignment is something that you really need so let's have a look it's finished the word alignment already

you know see some preliminary remarks they concern the application of the oh portal it's an application that runs in a browser window so as long as you don't close the door you will always have access to the data that you've been working on which is quite nice if something goes wrong you know the data is still there ok now we need to add some files so I will go to the operating system and select some files he has won by an English general and you can see now you can see the page of the oh portal they are four areas where is the top area where you can see a number of symbols they show what's going on at the moment they give you access to the help system they show you some statistics nice graphs showing how much time was spent doing as are and how much time was spent doing manual transcription if you find anything in the software if anything doesn't work just click on the feedback button and send us a message the little girl tells you that something that you can switch on desktop notifications some processes may take a while so if you don't want to sit and stare at the screen but do something similar and do something more interesting than just click on the bill and it will the system will automatically notify you the things changed finally you have the spartan phone settings then there's the second roll with three blue buttons and five verify and start processing the middle area is reserved for messages so there you can see what's happening or what's going on at the moment in this case it says one of your fire waiting to be verified then you'll see the column headings the column headers file upload speech recognition manual transcription word alignment the default speech recognition manual transcription and word alignment in phonetic detail are checked this means that our processing chain will go all the way from uploading through speech recognition you will do some manual transcription of correction and then finally there will be a word alignment and even phonetic details you might have seemed there is a little note that came up and speech recognition actually has finished now this is some of the file has arrived at the processing stage manual transcription and the symbol doesn't change to a little cog wheel instead has this little edit symbol let's go one step back first if you click on the icon on the on the green check Mark next to the speech recognition or in the speech recognition column you can see that there are a number of options available already you can have the results of the speech recognition in one of the formats presented here so par is the bus the Bavarian Archive for Speech Signals Partiturfile CTM is a format that's quite common for machine processing for speech recognition annot-json is new format that allows you to link the different elements on different annotation tiers which is very useful for later on processing TextGrid is a format might know when you've been working with Praat it shows the data in different tiers and on every tier all the items are marked with the precise time information table is a format that is very well suitable for R or Excel or any kind of statistics package and finally text is the raw text if we click on this we can see what Google made out of the audio file ok so this way you go through your file it might take a while until you've corrected everything Google made out of the audio quality isn't too good although we can and understand it very well perhaps the topic that the speech recognition and engine is not familiar with we don't know that's one of the problems we access external services we have no idea what they actually offer and they only provide a very limited set of configuration options with which we could adapt the services to our needs ok I will close this and you can see the whole transcript let's imagine that the transcript was ok and we're done with the transcripts so all we do now is press on Send transcription meaning that now we have finished the manual transcription it is sent to our server to the bus service and I'll word alignment starts again this takes a few minutes and if you really have a very true manual transcript of what was said word alignment should give you very good results for the actual alignment of the word it will give you more it will give you the alignment of the individual sounds but maybe this is something you don't need for your kind of research but I think the word alignment is something that really need so let's have a look at finished the word alignment

Manual: 4641 chars
Google ASR: 4604 chars
Levenshtein distance: 248
~ 5.3 % error rate

Manual transcript vs. Google ASR (called from OH Portal 22.02.2020)

ASR performance depends on audio quality and content

🔊 Technical signal quality

- 🔊 sample rate
- 🔊 quantisation
- 🔊 channels
- 🔊 compression

🔊 Recommendations

- 🔊 44.1 or 48 kHz sample rate
- 🔊 PCM quantisation
- 🔊 1 channel per microphone
- 🔊 .wav format

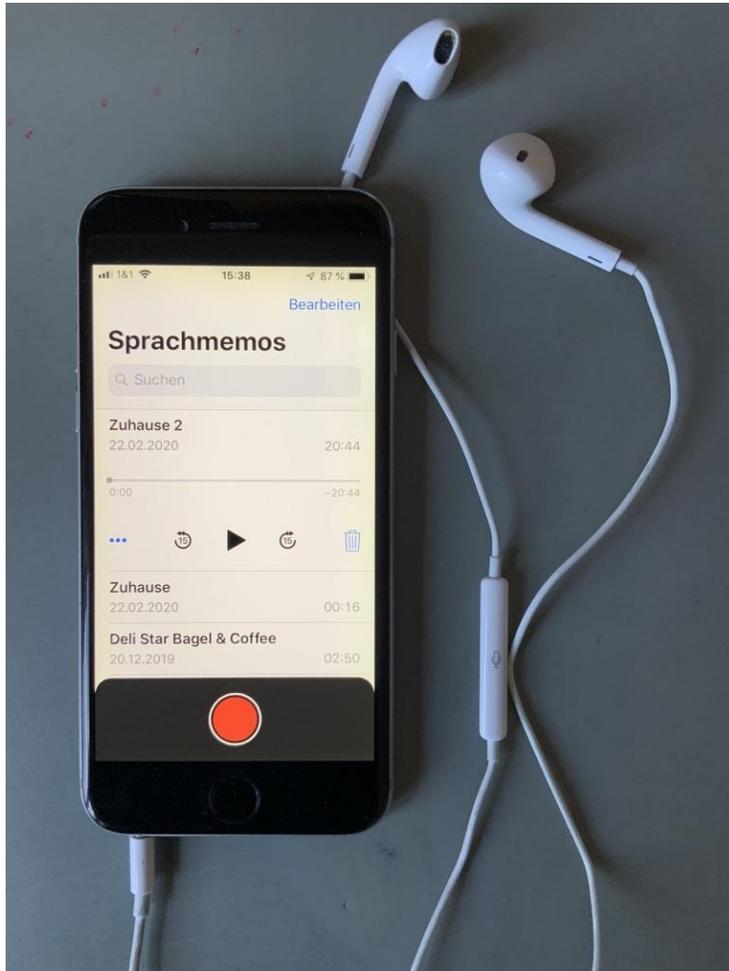
• Content

- 🔊 topic
- 🔊 speaking style & situation
- 🔊 environment noise
- 🔊 language, dialect, disorders

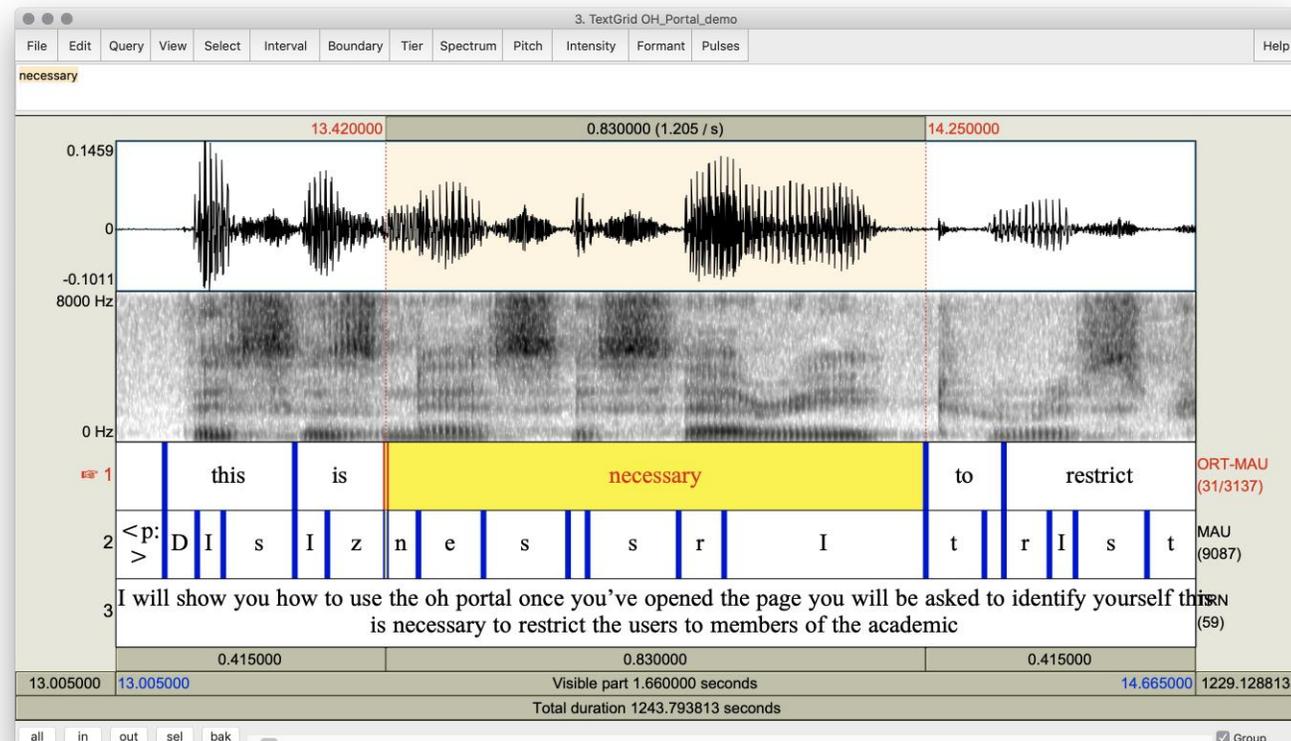
• Recommendations

- 🔊 aim for maximum speech & signal quality
- 🔊 close-talk or lapel microphones whenever possible

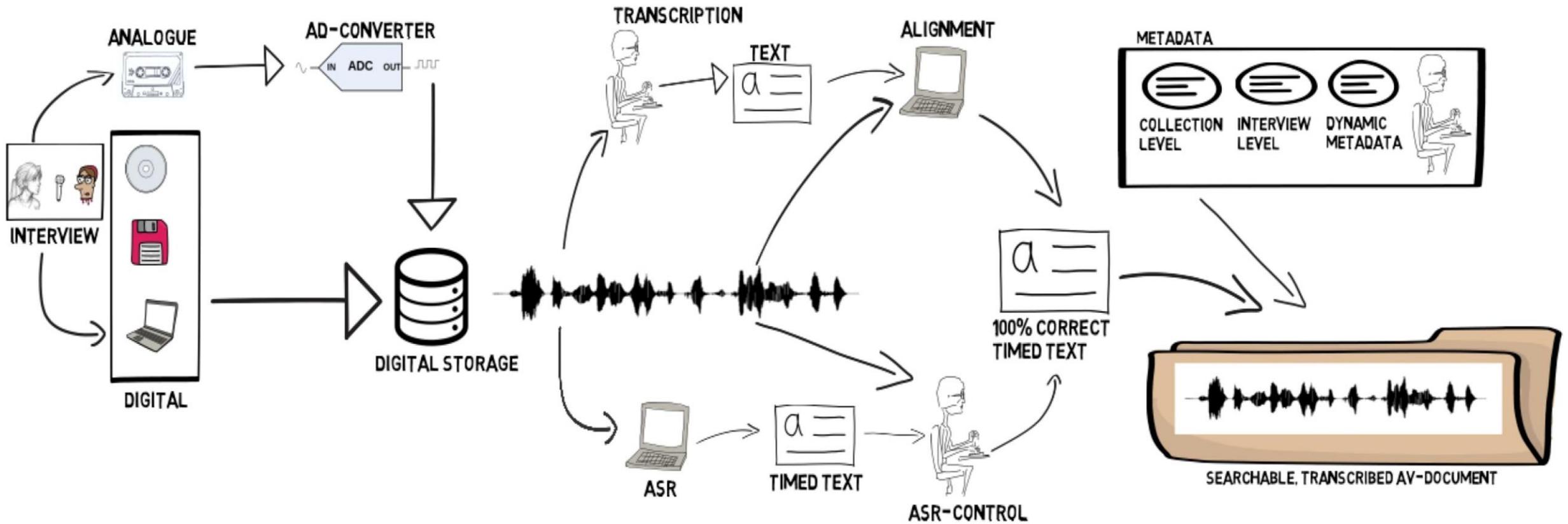
Less may be more...



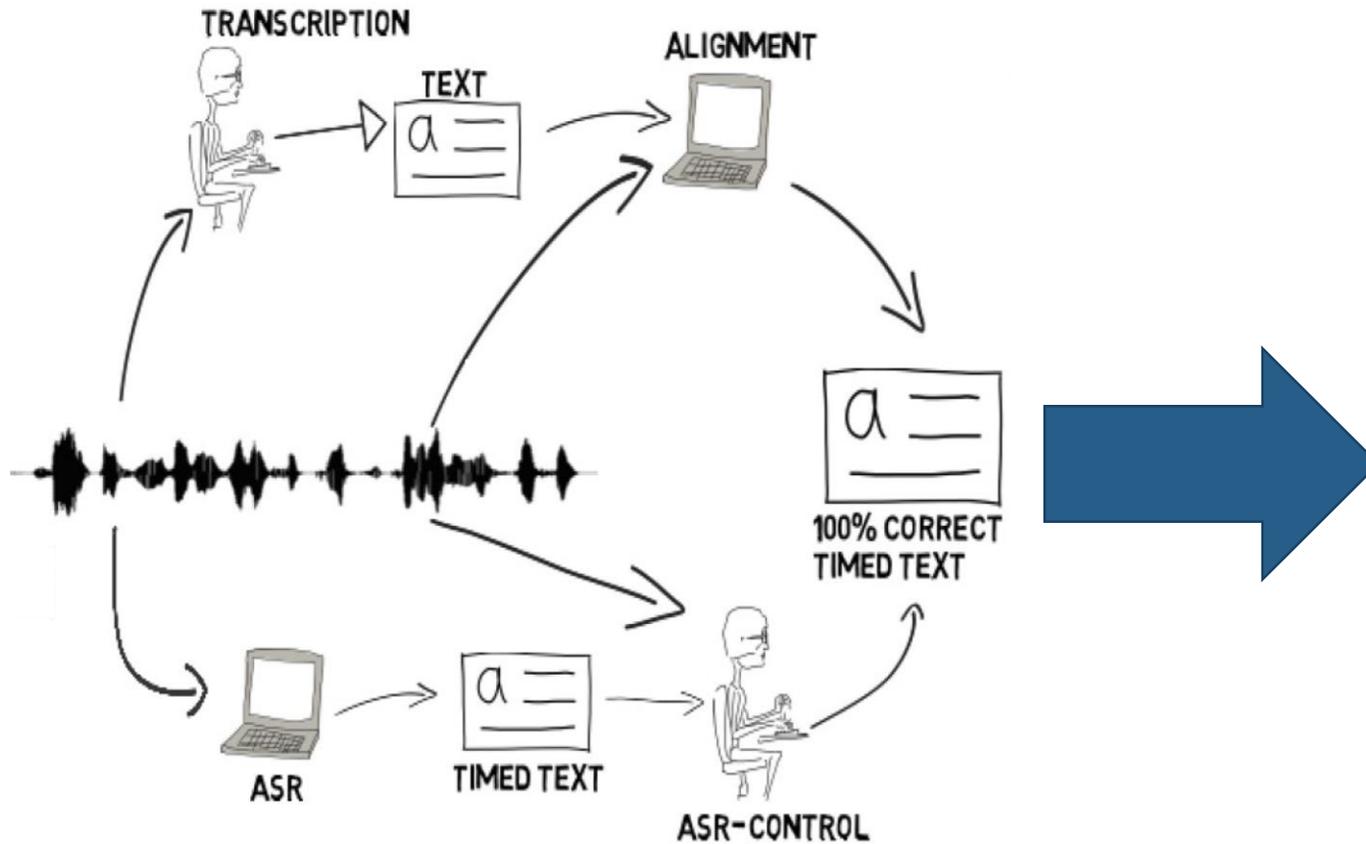
OH_Portal_demo.wav was recorded using an iPhone 6 with the standard headset microphone in a private living room



T-Chain Overview



Automating the T-Chain



1.
Add files

2.
Verify new files

3.
Start Processing

OH Portal

- 🌀 Graphical interface to state-of-the-art speech processing tools
 - 🌀 speech recognition
 - 🌀 manual transcription
 - 🌀 automatic word-alignment
- 🌀 Preconfigured workflow
 - 🌀 numerous export formats for further analysis
- 🌀 Several languages
 - 🌀 English, German, Dutch, Italian
 - 🌀 others in the pipeline...
- 🌀 Free for academic users
 - 🌀 requires authentication



OH Portal: Recommendations

🔊 Currently, the OH-Portal has some limitations

🔊 maximum file size approx. 200 MB

🔊 .wav audio file format

🔊 Thus, we recommend to

🔊 split your audio file into meaningful units (5-10 mins), e.g. introduction, housing situation, life at home, work, leisure time, etc.

🔊 use e.g. Audacity to

🔊 split stereo files to two mono files

🔊 downsample to 16 kHz

🔊 convert format to .wav

🔊 Don't use special characters in your file names!



OH-Portal v1.0.2

Ready

3. START PROCESSING

File Upload Speech Recognition Manual Transcription Word alignment Phonetic detail

Quickstart

- ### 1. Add files

You can drag & drop files (.wav & transcripts) 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. Verify new files

Before the new files can be processed it's required to set few options. Click on the "2. Verify" Button.
- ### 3. Start Processing

After all files are verified you can click on "START PROCESSING". The application starts the processing of pending tasks.



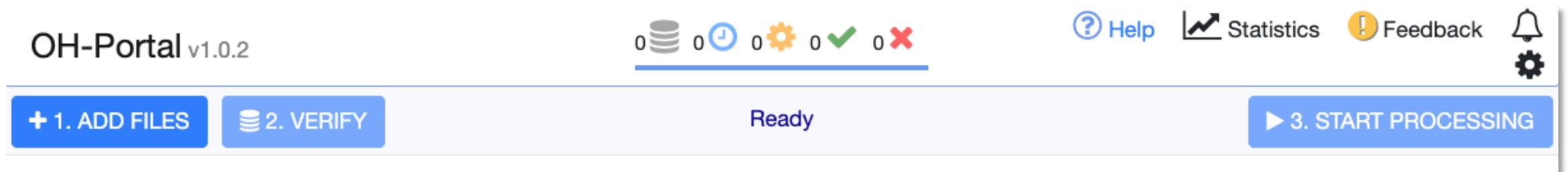
Using the OH Portal

🌀 Top row: version number, status icons, etc.

🌀 please use the feedback button to report problems

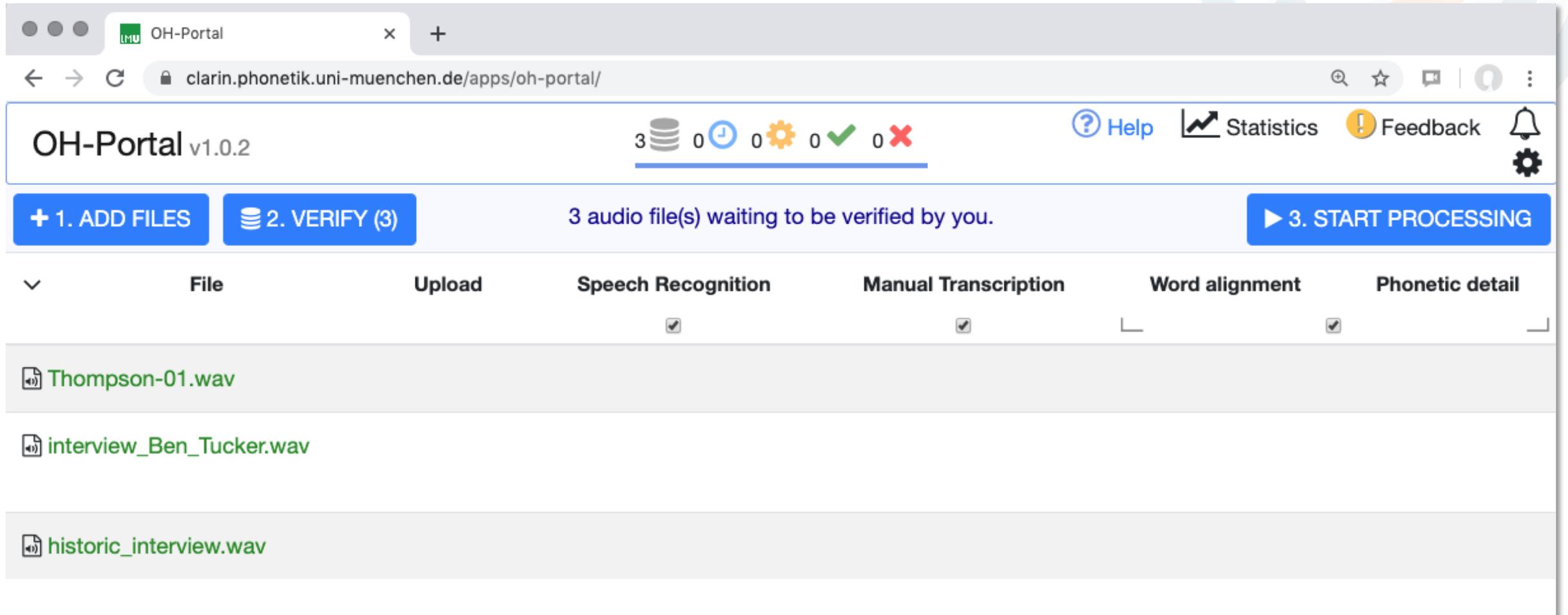
🌀 Action buttons

🌀 start and stop processing



The screenshot shows the top navigation bar of the OH-Portal. On the left, it displays 'OH-Portal v1.0.2'. In the center, there are five status icons: a database icon, a clock icon, a gear icon, a green checkmark icon, and a red X icon, each with a '0' next to it. On the right, there are four utility icons: a question mark for 'Help', a line graph for 'Statistics', a yellow warning triangle for 'Feedback', and a bell for notifications. Below the navigation bar, there are three main action buttons: '+ 1. ADD FILES', '2. VERIFY', and '▶ 3. START PROCESSING'. The word 'Ready' is displayed in the center of the interface.

OH Portal: drag & drop audio files



The screenshot shows the OH-Portal web interface. The browser address bar displays `clarin.phonetik.uni-muenchen.de/apps/oh-portal/`. The page header includes the text "OH-Portal v1.0.2" and a progress indicator with icons for database (3), clock (0), gear (0), checkmark (0), and error (0). Navigation links for "Help", "Statistics", "Feedback", and a settings gear are visible. A main action bar contains three buttons: "+ 1. ADD FILES", "2. VERIFY (3)", and "3. START PROCESSING". Below this, a table lists the audio files with checkboxes for "Speech Recognition", "Manual Transcription", and "Phonetic detail".

File	Upload	Speech Recognition	Manual Transcription	Word alignment	Phonetic detail
 Thompson-01.wav		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
 interview_Ben_Tucker.wav					
 historic_interview.wav					



OH Portal: select ASR service

The following files are going to be processed one after the other. Please check if all options are set as you wish. While you are selecting the language you can click on the logos of the service providers for further information like data storage policy and terms and conditions.

When you click "OK" you agree with the terms and conditions of the selected (third-party) services and the files are marked for further processing.

Language: English (GB) [Google] ▼

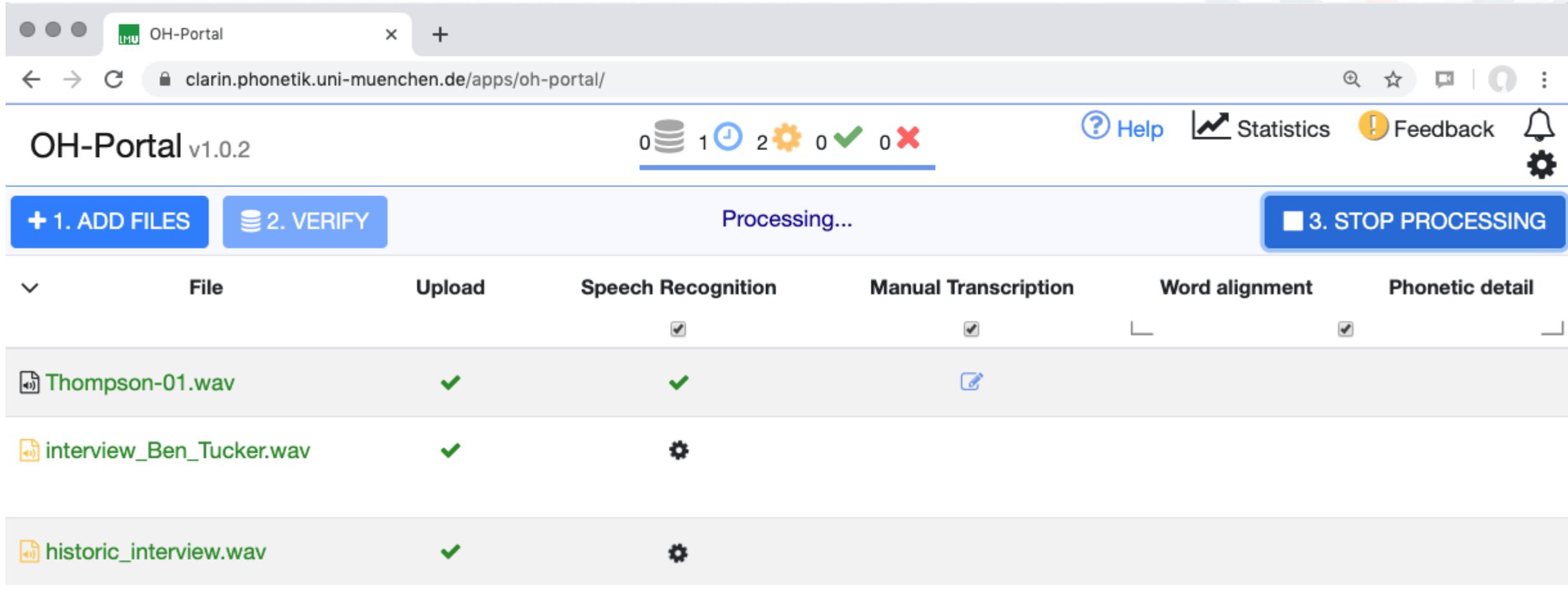
Language	Speech Recognition	Speech Recognition	Manual Transcription	Word alignment	Phonetic detail
German (DE) [Google]		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
German (DE) [EML]		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
English (GB) [Google]		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
English (GB) [Watson]		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
English (GB) [EML]		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
English (GB) [LSTEnglish]		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
English (GB) [Web]		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Italian (IT) [Google]		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Buttons: Cancel, OK

Click on logo to view terms of use

service providers may update services and terms of use without prior notice!

OH Portal: monitor progress



The screenshot shows the OH-Portal v1.0.2 interface. At the top, there is a navigation bar with a logo, version number, and a progress indicator showing 0 files, 1 clock, 2 gears, 0 green checks, and 0 red crosses. To the right are links for Help, Statistics, and Feedback, along with a notification bell and settings gear. Below the navigation bar are three main buttons: '+ 1. ADD FILES', '2. VERIFY', and '3. STOP PROCESSING'. The 'Processing...' status is displayed in the center. The main content area is a table with columns for File, Upload, Speech Recognition, Manual Transcription, Word alignment, and Phonetic detail. Three files are listed: 'Thompson-01.wav', 'interview_Ben_Tucker.wav', and 'historic_interview.wav'. The first file has green checkmarks in the Upload and Speech Recognition columns and a blue pencil icon in the Manual Transcription column. The other two files have green checkmarks in the Upload column and gear icons in the Speech Recognition column.

File	Upload	Speech Recognition	Manual Transcription	Word alignment	Phonetic detail
 Thompson-01.wav	✓	✓		└	└
 interview_Ben_Tucker.wav	✓	⚙			
 historic_interview.wav	✓	⚙			

OH Portal: download options

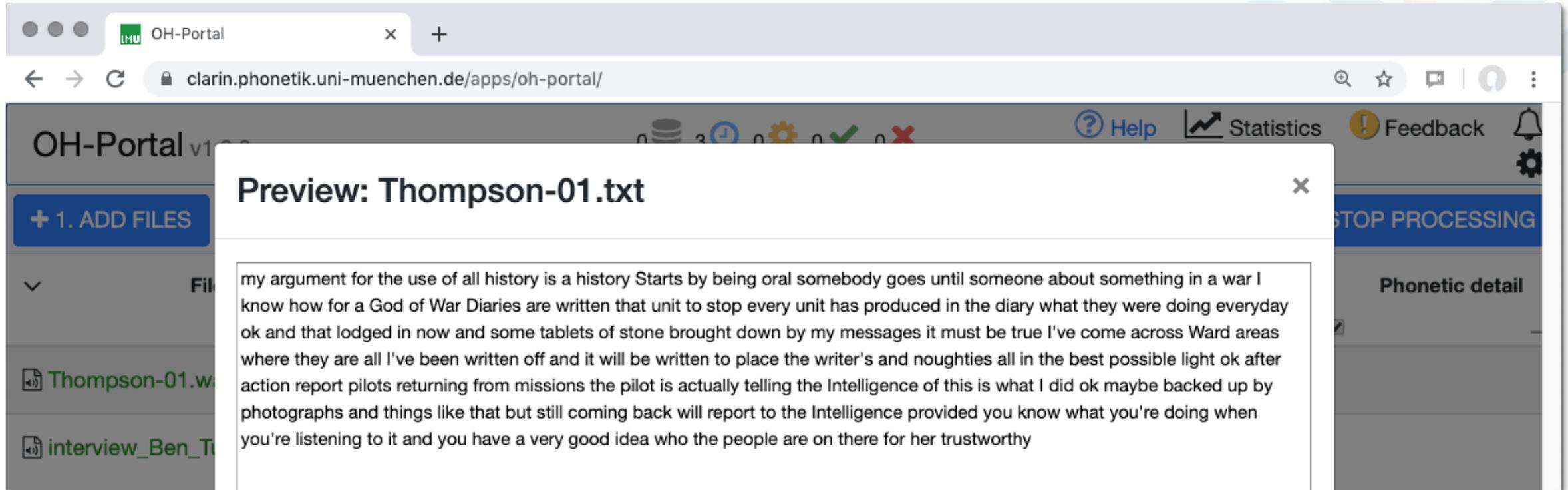
The screenshot shows the OH-Portal v1.0.2 interface. At the top, there are navigation buttons for '+ 1. ADD FILES' and '2. VERIFY', and a 'Processing...' status indicator. Below this is a table with columns: File, Upload, Speech Recognition, Manual Transcription, and Word alignment. Three files are listed: 'Thompson-01.wav', 'interview_Ben_Tucker.wav', and 'historic_interview.wav'. All files show a green checkmark in the 'Upload' column. The 'interview_Ben_Tucker.wav' row is highlighted, and a tooltip is open over it, showing a clock icon with the date '21.02.2020 15:05' and a timer icon with '00:01:06'. The tooltip also contains a table of conversion options.

Results	Conversions					
#	.par	*.ctm	*_annot.json	*.TextGrid	*.Table	*.txt
#1						

On Windows please use right-click and "Save as" to download a result



OH Portal: Preview result data



The screenshot shows a web browser window with the URL `clarin.phonetik.uni-muenchen.de/apps/oh-portal/`. The page title is "OH-Portal v1.0.0". The interface includes a navigation bar with "Help", "Statistics", and "Feedback" links. A modal window titled "Preview: Thompson-01.txt" is open, displaying the following text:

my argument for the use of all history is a history Starts by being oral somebody goes until someone about something in a war I know how for a God of War Diaries are written that unit to stop every unit has produced in the diary what they were doing everyday ok and that lodged in now and some tablets of stone brought down by my messages it must be true I've come across Ward areas where they are all I've been written off and it will be written to place the writer's and noughties all in the best possible light ok after action report pilots returning from missions the pilot is actually telling the Intelligence of this is what I did ok maybe backed up by photographs and things like that but still coming back will report to the Intelligence provided you know what you're doing when you're listening to it and you have a very good idea who the people are on there for her trustworthy



OH Portal: manual correction

built-in transcription editor Octra

- press tab to start and stop audio
- press return to open segment for transcription

OH-Portal v1.0.2

OCTRA: Thompson-01.par , Language: eng-GB , Audio duration: 02:04

File	UL	ASR	MT	WA	PD
Thompson-0...	✓	✓	⚙️		
interview_Be...	✓	✓	✍️		
historic_inter...	✓	✓	✍️		

ASR - Sprache : ENGLISH (GB) [GOOGLE] AKTION WÄHLEN ?

my argument for the use of all history is a history Starts by being oral somebody goes until someone about something in a war I know how for a God of War Diaries are written that unit to stop every unit has produced in the diary what they were doing everyday ok and that lodged in now and some tablets of stone brought down by

TRANSKRIPTION SENDEN ↗

OH Portal: transcription units

The screenshot displays the OH Portal v1.0.2 interface. The browser address bar shows the URL `clarin.phonetik.uni-muenchen.de/apps/oh-portal/`. The main header includes navigation links for Help, Statistics, and Feedback, along with status indicators (0 errors, 0 warnings, 0 successes). The current audio file is identified as "OCTRA: Thompson-01.par" with a duration of 02:04. A file list on the left shows three items: "Thompson-0...", "interview_Be...", and "historic_inter...". The "Thompson-0..." item is selected, showing a green checkmark for ASR and a gear icon for settings. The main content area displays the "OCTRA v1.4.1 (url)" interface, which includes a "SHORTCUTS" tab and an "OVERVIEW" view. A modal window is open, showing the ASR results for the selected audio. The modal includes a language selector set to "ENGLISH (GB) [GOOGLE]", a "SELECT ACTION" button, and a play/pause control. Below the control is a green waveform representing the audio. At the bottom of the modal, the transcribed text is displayed: "because somebody goes and tells someone about something and then they write it down".

OH Portal: quick check with overview editor

OCTRA v1.4.1 (url) TRN

Overview

20 20 0 20

No errors found

Transcript

Segment	Transcription	skip silence
# 01	my argument for the use of all history is a history Starts by being oral	▶
# 02	because somebody goes and tells someone about something and then they write it down	▶
# 03	yeah	▶
# 04	okay now	▶
# 05	back to to see this as so the auv writina it down making it wrong or the auv mav	▶

CLOSE

SEND TRANSCRIPTION ↗

OH Portal: finally, download transcripts

Download results by column ×

This creates a zip-archive of all the selected results and optionally conversions to other formats. If you do not select any additional conversions only the original results are added to the zip-archive.

Add conversions (optional):

<input type="checkbox"/>  CTM (.ctm)	<input type="checkbox"/>  BAS Partitur Format (.par)
<input type="checkbox"/>  AnnotJSON (_annot.json)	<input type="checkbox"/>  TextGrid (.TextGrid)
<input type="checkbox"/>  Table (.Table)	<input type="checkbox"/>  Plain Text (.txt)

[Get package](#)

[Close](#)

- 🌀 File-wise, by clicking on the check mark
- 🌀 Batch-wise, by clicking on the column head

Recommended workflow

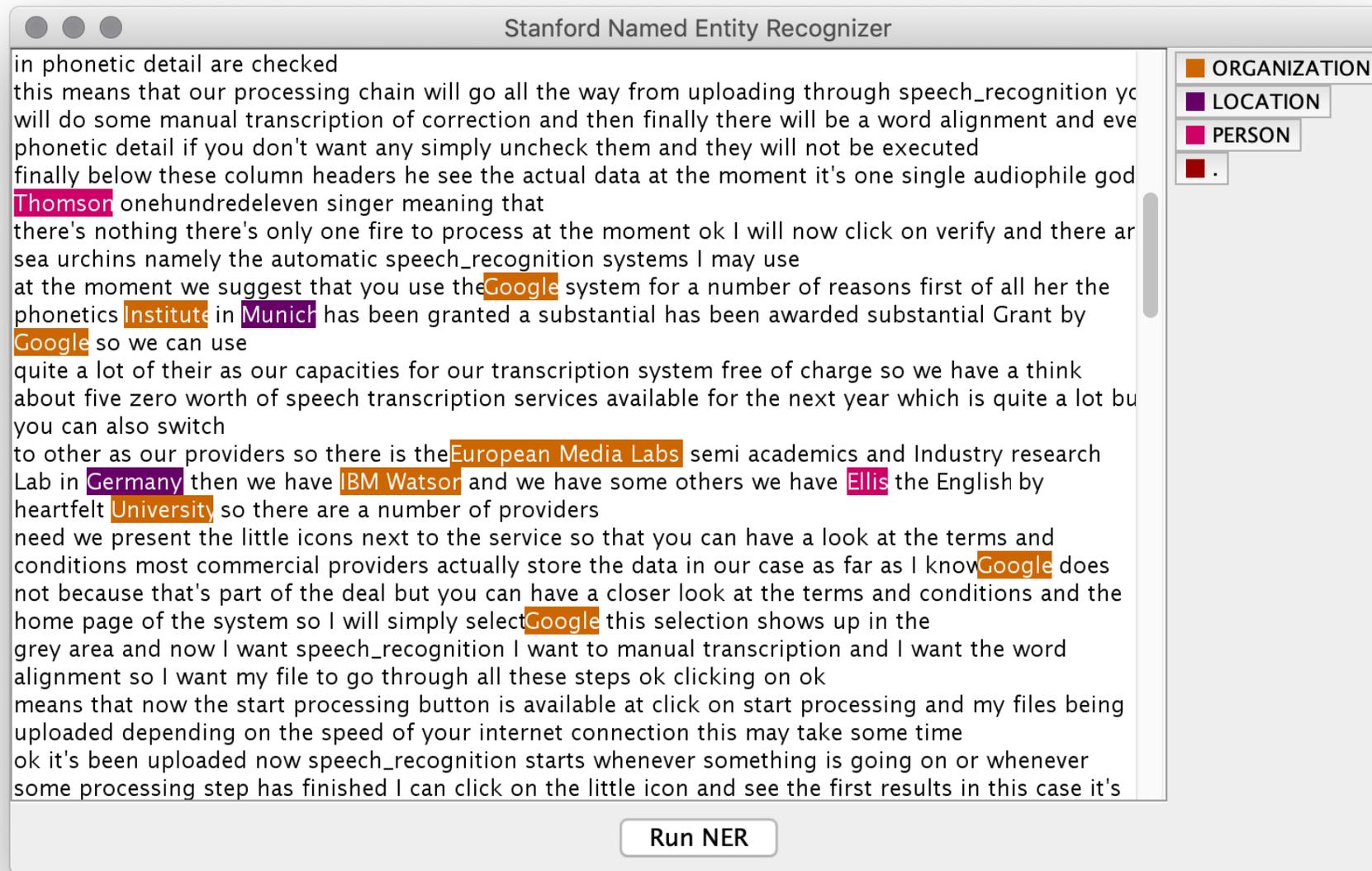
- 🌀 Split recording into meaningful units
- 🌀 Run ASR on recordings
- 🌀 Check ASR results
 - 🌀 👍 : skip manual correction
 - 🌀 👎 : do manual correction
- 🌀 Run word-alignment
- 🌀 Download transcript(s)



Where to now?



In-depth text analysis: Named Entity Recognition



Stanford Named Entity Recognizer

in phonetic detail are checked
this means that our processing chain will go all the way from uploading through speech_recognition you will do some manual transcription of correction and then finally there will be a word alignment and even phonetic detail if you don't want any simply uncheck them and they will not be executed
finally below these column headers he see the actual data at the moment it's one single audiophile god **Thomson** onehundredeleven singer meaning that
there's nothing there's only one fire to process at the moment ok I will now click on verify and there are sea urchins namely the automatic speech_recognition systems I may use
at the moment we suggest that you use the **Google** system for a number of reasons first of all her the phonetics **Institute** in **Munich** has been granted a substantial has been awarded substantial Grant by **Google** so we can use
quite a lot of their as our capacities for our transcription system free of charge so we have a think about five zero worth of speech transcription services available for the next year which is quite a lot but you can also switch
to other as our providers so there is the **European Media Labs** semi academics and Industry research Lab in **Germany** then we have **IBM Watson** and we have some others we have **Ellis** the English by heartfelt **University** so there are a number of providers
need we present the little icons next to the service so that you can have a look at the terms and conditions most commercial providers actually store the data in our case as far as I know **Google** does not because that's part of the deal but you can have a closer look at the terms and conditions and the home page of the system so I will simply select **Google** this selection shows up in the grey area and now I want speech_recognition I want to manual transcription and I want the word alignment so I want my file to go through all these steps ok clicking on ok
means that now the start processing button is available at click on start processing and my files being uploaded depending on the speed of your internet connection this may take some time
ok it's been uploaded now speech_recognition starts whenever something is going on or whenever some processing step has finished I can click on the little icon and see the first results in this case it's

Legend:

- ORGANIZATION
- LOCATION
- PERSON
- .

Run NER

Screenshot of Stanford
Named Entity Recognizer (v.
3.9.2)

Used with the included basic
3 class NER tagger trained
on standard text corpora

In-depth text analysis (e. g. voyant-tools.org)

The screenshot displays the Voyant Tools interface with the following components:

- Word Cloud:** A word cloud on the left side, with 'speech_recognition' and 'alignment' being the most prominent words.
- Text Viewer:** A central text area showing a document snippet. The word 'speech_recognition' is highlighted in yellow in the original image.
- Line Graph:** A line graph on the right showing 'Relative Frequencies' across 'Document Segments (1 to 10)'. The graph tracks five categories: alignment (purple), ok (blue), speech_recognition (green), transcription (cyan), and word (magenta).
- Context Table:** A table at the bottom right showing context for the highlighted term 'speech_recognition'. It lists document segments and the surrounding text on both sides.

Context Table Data:

Segment	Left	Term	Right
1)	the column headers file upload	speech_recognition	manual transcription word alignment the
1)	transcription word alignment th...	speech_recognition	manual transcription and word alignment
1)	the way from uploading through	speech_recognition	you will do some manual
1)	sea urchins namely the automatic	speech_recognition	systems I may use at
1)	area and now I want	speech_recognition	I want to manual transcription
1)	ok it's been uploaded now	speech_recognition	starts whenever something is going
1)	so now waiting for the	speech_recognition	the overall duration of the
1)	from our server to the	speech_recognition	engine at Google this takes



UPCOMING EVENTS AND WORKSHOPS

- SSHOC Consortium Meeting, March 24, 2020, Utrecht: **SSHOC Workshop - Linking Social Survey and Linguistic Infrastructures through EOSC**
 - <https://www.sshopencloud.eu/sshoc-workshop-linking-social-survey-linguistic-infrastructures-eosc>
 - First SSHOC ASR Use Case on panel interview data, pilot
 - (CLARIN) tools for analysis of audio
- **LR4SSHOC Workshop @LREC, May 11, 2020, Marseille:**
 - <https://www.clarin.eu/LR4SSHOC>
 - Cross disciplinary tools for analysis of interview data
- **ICMI Workshop, 25-29 October 2020, Utrecht: Speech, Voice, Text, and Meaning**
 - <http://icmi.acm.org/2020/>
 - A Multidisciplinary Approach to Interview Data through the Use of Digital Tools

Thank you for your attention!

Questions?

Please put them in the chat box.

Slides and a recording will be sent to all registered delegates.

Join our community



<https://www.sshopencloud.eu>



@SSHOpenCloud



info@shopencloud.eu



[/in/shopencloud](https://www.linkedin.com/company/shopencloud)

