Thinkie 3 for Thinking Aloud of 2021

https://apps.apple.com/de/app/thinkie-3/id1552765360

BRIGITTE ENDRES-NIGGEMEYER

https://www.noapps.org , brigitteen@gmail.com

Summary

Thinkie 3 supports Thinking Aloud on iPhone and iPad.

Launching Thinkie in 2021 one asks in which environment the app is being exposed. The article starts with a description of the current technical and methodological state of Thinking Aloud (TA). Since the beginnings in 1980, things have changed.

After that, attention shifts to Thinkie 3 itself. As a mobile IOS application Thinkie 3 stores its data in a cloud container of the user. The data structure includes metadata, audio and video files, audio clips and transcripts. The server-based Apple Speech-To-Text helps with transcription. Users are supported in data capture and initial data analysis. Files can be exported for further handling.

For users there are guides in English and German.

As concurrent approaches are missing in IOS, a Thinkie 4 with new/improved features might be an option.

Keywords

iPhone, iPad, iCloud, IOS, Thinking Aloud, methods, data capture, audio, video, NLP, speech, Speech-To-Text, transcription, verbal protocol, automatization, user interface, UX, application areas

Overview

The Thinking Aloud (TA) method is not new. On can study its basics in Ericsson & Simon 1980 or later and learn how to do it from van Someren et al. 1994.

But what is its state of today? This question popped up when Thinkie 3 was launching for the AppStore and for users, after a long development story. In which environment does the Thinkie app drop in? What will the app contribute there? And which perspectives exist there for future development?

Since Ericsson & Simon 1980 Thinking Aloud has proved its basic value, but it has enlarged and seized the facilities of its context.

It adapted to the current state of technology:
- Today audio and video recording is the standard, eye tracking may join.
- Internet and WWW with their applications are available from everywhere.
- Speech recognition made big steps toward everyday applicability.

TA methods expanded according to the needs of different application areas:
- In order to follow the thinking process TA often combines with other techniques.
- In many disciplines substantial data models have been set up. They store empirical results of TA according to the demands of the current study.
- Transcription and data analysis widen far beyond what Ericsson & Simon imagined at the beginnings.
- TA is made economical. Especially in Usability Engineering verbal protocols of TA are partially exploited as far as needed only. Transcription and interpretation is reduced or avoided, eg. to observer notes for further discussion.
- The interaction of a user / test subject with the things around is recorded and therefore more and more a verifiable part of the thinking process.
- This would lead to more videotaping and extending media from sound to multimedia.
- With user roles in the natural environment obtaining more attention data capture in the field is practised more in common.

After the discussion of the current TA context, Thinkie 3 itself is characterized:

- Thinkie 3 is a minimalistic and mobile app supporting TA data capture and transcription.
- Thinkie is available world-wide on IOS. It needs an iPhone and an iPad with iCloud connection.
- The sound of the spoken verbal protocol is captured on the iPhone, the video on the iPad.
- All data is stored in a private cloud container of the user within a custom data structure.
- All files can be exported to the user’s system environment.
- On the iPad the audio verbal protocol can be segmented for manual transcription or submitted piecewise to the server-based Apple Speech-To-Text API. The transcript can be corrected.
- Comparable other apps did not yet appear in IOS. Thinkie itself may improve eg. by integrating more developed tools for transcription and data interpretation. So that Thinkie 4 might be an option.

Thinking Aloud as seen today (2021)

The extensive basic guide for Thinking Aloud application still is Someren et al. 1994.

Working memory

We think using our brain and the memory system. By and large, the memory system model set up by Atkinson & Shiffrin 1968 is still in use today.

It defines an input area that stores and repeats sensory data for a short time.

The working memory (Short-term memory - STM) with its restricted capacity is the main place of thinking (Baddeley 2003). There, content items (chunks) are quickly overwritten by newly incoming items and forgotten. In long-term memory (LTM) information is stored in large sizes for long time. Content from STM is being stored there.

During thinking all the time material is fetched from long-term memory and put into the working memory. In addition sensory data is integrated into thinking processes.

The inner structure of the working memory and its connection with the long-term memory is illustrated by fig. 1. There are alternative theories about it (Cowan 2010, Quak et al. 2015). A more detailed description is found in Amin & Malik 2013.

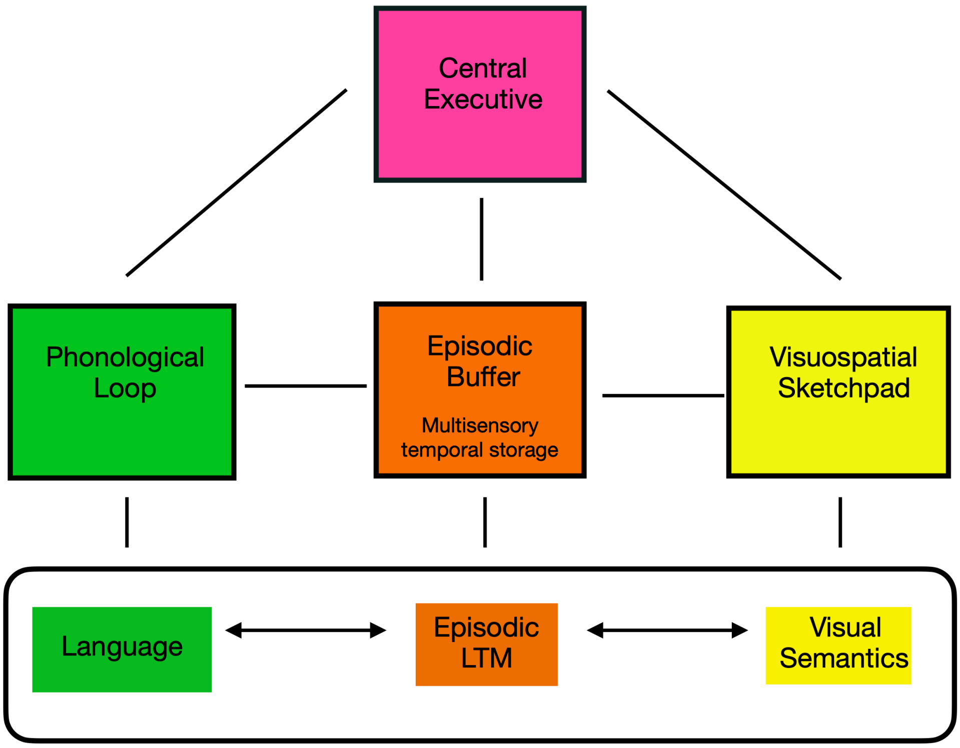

Fig. 1. Working memory according to Baddeley, adapted from Park & Gooding 2014

Think and verbalize: Thinking aloud

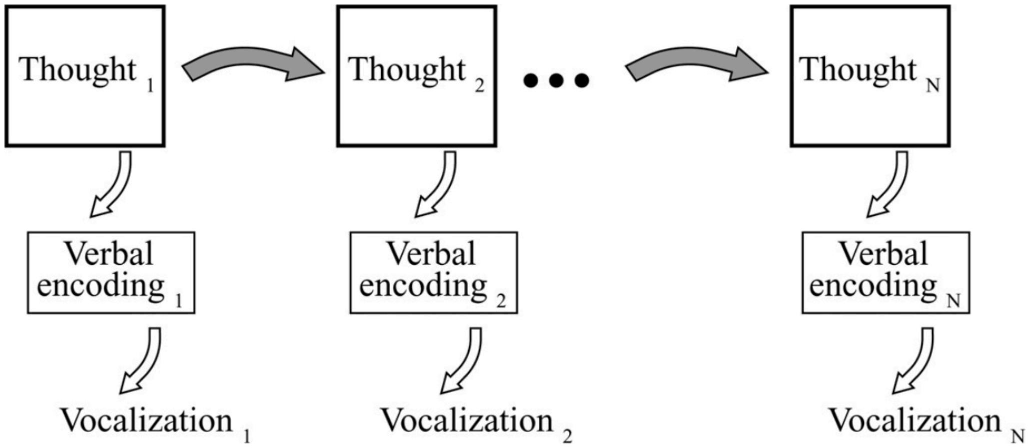

Fig. 2. Thinking process with verbalization of thoughts / thinking steps, from Ericsson 2006

When thinking follows a structured path, one thinking step - one thought - follows the other. When this happens with thinking aloud (with talking), the individual thinking steps are verbalized while they occur, so that they are accompanied by spoken utterances. This model of thinking and Thinking Aloud (TA) is due to Ericsson and Simon 1980. It is shown in fig. 2 taken from Ericsson 2006.

In everyday life talking while performing a task occurs frequently, in particular when a sequence of steps must be executed, such as setting a some switches before starting a machine.

When a person keeps talking (“verbalizing”) during a structured thinking effort, this is TA and results in a spoken verbal protocol. It maps the thinking process taken from the content of the working memory. The account is true in what it says, but not necessarily complete. It is data: „Verbal Data is Data“.

The rendering of the current working memory content may be incomplete in several respects:
- Thinking steps may be missing, in particular if they are too automated for showing up in awareness.
- Content is forgotten, in particular if it is overlayed quickly with new chunks.
- Speaking is slower than thinking. If talking is too slow to render all thoughts, some of them may fall by the wayside.

Because during a concurrent / synchronous verbal protocol thinking and talking happen in parallel the process tends to slow down, but it keeps its structure.

When a verbal protocol is captured during a study, almost always an observer is present (fig. 3). He or she reminds the test subject to keep talking in case of silence. Beyond this any manipulation of the test subject should be avoided according to the predominant opinion.

Fig. 3. Participants of a Usability Test / Thinking Aloud from Moran 2019

If it is not possible or not intended to record a verbal protocol, the observer may note important utterances.

So far what the theory says.

Boren & Ramey 2000 investigated the practical implementation of TA in Usability Engineering and found widely diverging ideas. Bernardini 2001 inspected the application of TA in Translation Studies and detected comparable method differences or shortcomings.

Hughes & Parkes 2003 report on the state in Software Engineering.

Thinking depends on context

Thinking is bound to its environment. Things around the person stimulate concepts or link the imagination to large-scale mental spaces:

- The stop sign at a crossing inspires how to go on in traffic: stop, check whether there is somebody with the right of way. If so wait to let him or her pass, if there is nobody, go ahead.
- A tango tune may remind Argentina or make you dance some steps.

In situations with TA it is all the same. A test subject who is to work on a user interface needs to navigate on it using buttons, the own voice or any other means. The same is true for any other objects, for instance a phone, a dictionary or a notepad. By their presence and use they inspire thinking.

Noise and other disturbances should be avoided during TA. Generally speaking who is comfortable thinks better.

Today almost always a video documents how the surroundings look and sound like during a TA process.

A video displays the physical behavior of the test subject. The objects in the environment that the personn considers or moves - the “affordances” (Gibson 1986) - hint at what her or his current thinking is about.

From the verbal protocol and the video recording one can reconstruct what the person was thinking and doing (triangulation). This improves the quality and the accuracy of interpretation.

Lab studies and field studies

Very frequently TA studies are done in the lab, the place where the researchers are working and the technical infrastructure is available and well-known. Financial reasons may exclude other solutions, especially in the industry,

In Usability Engineering lab studies are the normal approach (Nielsen 1994, Fan et al. 2020b, McDonald et al. 2012). Some loss of validity is accepted (Kieldkov et al. 2004, Joe et al. 2015). One finds most use problems, but not all. Field studies are appropriate if they are really needed (Kjeldkov & Skov 2014).

Field studies are required if capturing data is not possible outside their normal environment, for instance at the patient bedside (Aitken & Mardegan 2000, Han et al. 2007), in cancer aftercare (Jaspers et al. 2009), during car driving (Monsalve et al. 2020), during a virtual-reality earthquake (Feng et al. 2020), or at a designer’s studio (Pringle & Sowden 2017).

For exploiting verbal protocols in detail the environmental features should fit. In the natural surroundings thinking processes run better and meet reality better than possibly biased performances in a lab. Where this is needed or desired one accepts the additional effort of a field study in terms of costs, technical outfit and traveling and works with the test persons in their home environment.

Concurrent / synchronous oder retrospektive?

In addition or alternatively to a concurrent / synchronous verbal protocol one can ask for a retrospective protocol. For this, the test subject has to fetch data from long-term memory. Gaps may occur and interpretations may have shifted.

Often one asks the test subject for a separate interpretation of the experience, orally or using a questionnaire. Talking the verbal protocol over with the test subject is just as usual. This can also serve to explain doubtful passages of the protocol.

A comparison of the procedures is given by Alhadreti & Mayhew 2018. They advocate the less time-consuming concurrent method. Olmsted-Hawala et al. 2010 provide an earlier discussion of different capture modalities.

Data models

After data capture interpretation and data analysis is on. Ericsson & Simon 1980 do not mention that very much. On fig. 2, one thought follows the other. What the thoughts mean and contain remains hidden.

If somebody states “I do not know how to get on” and somebody else says “To the right, let’s go!”, both utterances differ in their cognitive content and in the actions they accompany and trigger. Their properties make them join different classes. The first example may go to metacognition or self-monitoring, the second to “starting action”. Besides this, the first example contains only one item, stating a lack of knowledge, whereas the second one expresses two thinking acts: the activation of situational knowledge (“to the right”) and the onset of an action (“let’s go!”).

Thoughts / thinking steps can be organized depending on their properties. Inside they may be composed. A thinking act may integrate a structure of several items.

In a TA study thinking acts are modeled to fit the aims of the research, so that the data model is justified by the theory of the application area.

When dealing with a user interface in Usability Engineering, the controls (start button, exit etc.) are an obvious initial classification criterion for thinking acts. When users experience a problem with a control, one can ask for the time needed to solve it or its reason.

Models may be preset, but they may as well be developed from the observed data. In Endres-Niggemeyer 1998 cognitive summarization agents have come up first from a selftest of the author and in the following from working processes of six colleagues observed with TA. The methods of data modelling came from Knowledge Engineering (Schreiber et al. 1993, Fensel et al. 1998).

Endres-Niggemeyer 2000 describes a demo implementation displaying the summarization agents of different functional groups as types of insects, for instance with bees doing information collection.

In order to decide whether Facebook news are correct or wrong, Freilich 2019 discovers in her test subjects three main strategies:
- searching for more
- knowledge carries the most weight
- every detail needs to fit

The strategies were in part taken over from other models in a deductive approach, but in in part also developed inductively from the observed data.

The verbal protocols also showed that many opinions in social networks cannot be classified as right or wrong. The subjects used secondary features for their decision, eg. correct quotes.

Zhou & Lin 2012 use a strategy model to find out how students translate texts from English to Chinese and vice versa. They adopt their model from the literature and expand it for their own research.

Transcription ?

Whether one should transcribe a spoken verbal protocol is a matter of debate. In any case transcribing takes time. How much transcription is needed depends very much on the target of the study. Because the transcription includes verification and rendering of the audio content, it implies an interpretation / a beginning analysis of the data.

Following Nielsen 1994 and later, one skips the labor-consuming transcription in Usability Engineering. With a „simplified Thinking Aloud“ an observer notes the problems which the test person encounters and verbalizes when checking a user interface. These notes suffice for further processing.

With their „Instant Data Analysis (IDA)“ Kjeldskov et al. 2004 expand the Nielsen approach. A tutor and an observer in a different room watch the TA session of the test subject. At the end, all participants combine their impressions in an idea generation process („Instant Data Analysis“). Because all participants cooperate, most weak points of the interface are captured. They are ranked on their severity.

Joe et al. 2015 apply IDA inspecting the use of a medical platform for elder patients. They obtain satisfactory results as well with having the test subjects use the platform while doing TA and after that have them participate in an idea generation session with pin board and cards. The group mentions the problems and orders them by consensus.

If the aim is to check how persons behave in specific situations (their sentiments, possible disorders, wrong taps, etc.) , transcribing and analyzing the referencing partial audio sequences only will do the job. The start and end times of the spans are noted in the transcript.

Feng et al. 2020 investigate how people behave in an immersive Virtual Reality (VR) earthquake crisis in a hospital in Auckland. During the event the verbal protocols are captured and synchronized with the video on the VR glasses. Both tracks are stored in an MP4 file.
A simple coding scheme specifies the expected situations, such as during the earthquake or after. The situations are equipped with typed statements about what the test persons do or how they check the situation, eg. running down the scales following others or testing whether the elevator works. Human raters tag occurrences of these situations with the situation and one of its typed behaviors.

Sakar & Chakrabarti 2007 interviewed colleagues with much transcription experience on problems they met frequently. They developed technical support for handling these problems. Their tool set for supporting manual transcription halved the time spent for a transcription.

Analysis tools like ODCS (MacLin & MacLin 2005) and CAPAS (Crutcher 2007) or commercial tools such as MAXQDA and many others accept different inputs such as audio, video, written verbal protocols or numerical data.

[2]

A (m)	okay so.	((ea)) Okay	((1s)) legen
[KO]	*Geräusch im Raum, 0,5s*	*Geräusch im Raum, 0,2s*	

[3]

A (m)	Sie los. Sprechen Sie quasi ganz frei und unbefangen, ja?		
M (m)		((lacht))	Ja, also
M (m) [k]			*lachend*
[KO]		*Möbelrücken*	*Möbelrücken*

[4]

M (m)	es is so/ es geht • darum, dass ähm… Äh wir sind verheiratet, haben n
M (m) [k]	

[5]

A (m)		Hm̌
M (m)	gemeinsames Kind ((ea))	Und ich hab aus erster Ehe • • ne Tochter,

[6]

A (m)		Hmhm̌	
M (m)	Bin geschieden,	Und meine Tocher is	Endee zweitausendund
M (m) [k]			*langsamer*

Fig. 4. Transcript in score format: sample with a track for the lawyer, the client and the background, taken from Pick 2015

Transcribing the audio verbal protocol means often more than a simple transposition from spoken to written text. An example is given by Hutchison et al. 2015. They code pauses and gestures (in their case of information seeking) in order to grasp the cognitive process better.

What the test person said is not that always easy to understand. For noting it, one may need to listen several times and to interpret, sometimes to no end so that one has to accept and note an unclear spot. The transcriber conforms to the orthographic rules as far as possible although they are a feature of written, not of spoken text.

The wording is important, but it is not all. Even in a simple transcription one notes pauses, interjections („ahh“, „hmm“), changes of pitch or volume, expressions of emotional states like laughing or sighing. The help to separate steps of thinking and action.

For more sophisticated transcripts the verbal protocol may be noted as a score with tracks for all sound sources. In fig. 4 a lawyer, the client, and the background are noted in parallel. For both lawyer and client the track adds a speaking mode feature (“laughing”, “slower” in the sample). All tracks of the written verbal protocol can store additional properties, for instance their start and end times or classification codes. In the sample, the environment has a track of its own, containing “moving furniture”.

Guides for manual transcription on the web often refer to different applications and system environments: Klein 2015, Setting et al. 1998, Rincon 2018, Dresing et al. 2015, or look at platforms and tools like media-ai or audiotranscription.de where automatic transcription is offered, too.

When an automatic transcription is possible one will integrate it, normally expecting that the delivered text needs some correction. Features like pitch or background noises can be entered by hand.

External providers like NVIVO , Amberscript or Otter can take over the whole transcription task.

Data analysis

Data analysis means to check a set of verbal protocols for data values. Analysis conforms to the conditions of the basic structure and the target object of the study.

A study may integrate more input channels than just verbal protocols possibly accompanied by a video and some eye-tracking results. Other input may come as questionnaire, as retrospective interview with the test subject, or as a cost analysis. They have to be included into the overall result.

The overall study aim and result may be restricted, eg. if it just records simple values like the time intervals needed for the user's web site handling. It may however ask for much more data supply, eg. if it asks for the sites the user accessed with which success for answering his or her question.

A quantitative statistical analysis targets numerical values like the time or the number of trials. As spoken or written verbal protocols are not of numerical but of text type, they will most of the time need a qualitative analysis. It may be combined with a quantitative investigation.

A deductive approach would seek the data for instances of preset classes while an inductive method abstracts from and aggregates categories discovered in the data. Both approaches may mix.

Depending on the target of the study methods can vary in a wide area. Some examples:

- Creative thinking for garden design (Pringle & Sowden 2017)
- Rules for autonomous driving (Monsalve et al. 2020)
- Decision-making during an earthquake in a hospital (Feng et al. 2020)
- Interpretation of radiographies (Yoon et al. 2020)
- Translation (critical overview by Bernardini 2001)
- Methods during a translation process (Sun 2011)
- Learning to read and to understand texts (Wang 2016)
- Understanding of complex texts (Dahl et al. 2021)
- Foreign language tests (Green 2007)
- Intercultural differences in usability testing of Chinese and Danish users (Shi 2010)
- App for enhancing physical movement (Fischer et al. 2019)
- Misinformation in medical information seeking (Ghenai et al. 2019)
- Self-assessment of patients (Al-Janabi et al. 2013)
- Nursing in intensive care (Han et al. 2007)

Methods of qualitative data analysis are provided by NN 2014?, 2019, Lester et al. 2020, with reference to support tools by Gizzi & Rädiker 2021, for MaxQDa and for NVIVO by Jackson & Bazeley 2019. PAT Research reviews available systems for qualitative data analysis. They inform about other providers, too.

Elzakker et al. 2008 propose combined methods for qualitative and quantitative investigation of mobile maps use in the field and in the lab. Their obtain the best experiences with a combination of video recordings, screen-logging, synchronous TA and semistructured interviews of the users.

Software tools for data analysis

With more inputs coming in, data analysis may become more difficult. The test subject may talk (sound input), she may move an object or take a note (video input), she may change the viewing direction and attend to some point not considered before (eye tracking - Oh et al. 2013).

Levy & Ransdell 1994 observed how texts are written. For this purpose they recorded the spoken verbal protocol and the keystrokes, mapping both on a video. The video was displayed on a TV set. Raters coded it adding start and end times and frequency of the sequences. Like this the main subprocesses of text production (planning, reviewing, generating text, revising according to Kellogg 1987) could be discovered in the data.

ELAN (fig. 5 - Trippas et al. 2017) is a tool for working with multitrack input. It supports segmenting intervals from video and audio and annotating them. Additional helps tools are available. The annotations may contain a whole transcript.

Asselin & Moayeri 2010 work on a market research system for use analysis (Morae) with spoken verbal protocols and video recording. Their aim is learning more about two youngsters' skills on the web. The youth did their homework at their home computers. Both the video and the verbal protocol data was exploited for their cognitive and emotional reactions.

Maher et al. 2018 describe how they display design processes combining NVIVO and sticky notes. Thinking processes can be rendered by voice (TA), but also haptically. They can be tagged on both channels. Fig. 6 shows their combined output from NVIVO data and sticky notes.

Ghenai et al. 2019 observe how web browser information presentation influences which messages users believe to be right or wrong, They record a video with sound track and tag the vision direction of their test users. The verbal protocols are taken synchronically and retrospectively. During the retrospective recording the eye tracking results helped to remind the respective moment. The data was coded with NVIVO software.

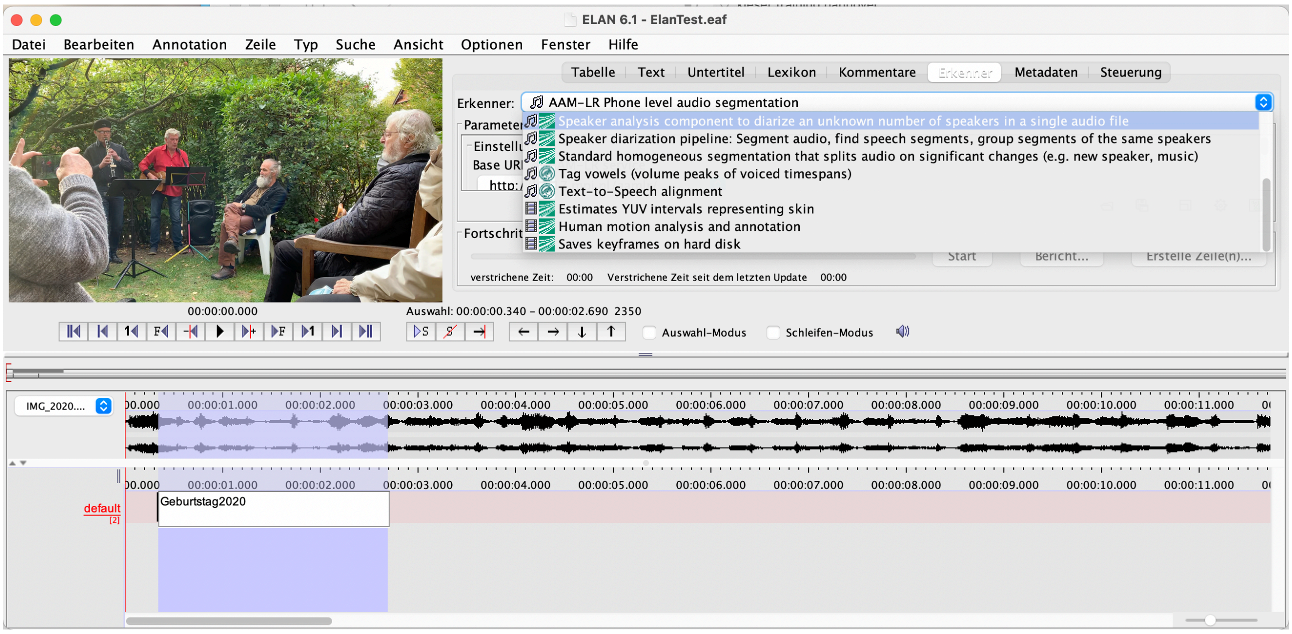

Fig. 5. ELAN: Video and audio tracks, annotation (“Geburtstag 2020”) and helps for recognition

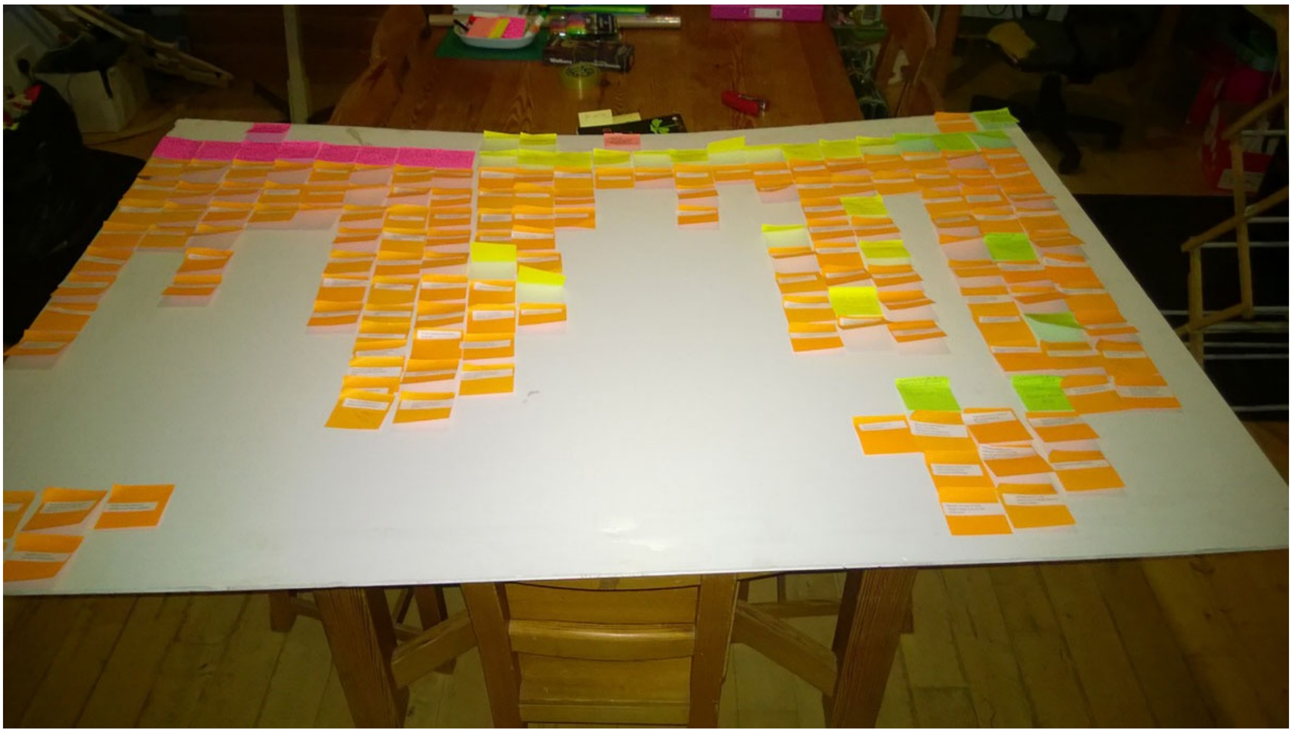

Fig. 6. NVIVO-Output with sticky notes: combined display by Maher et al. 2018

Thinking aloud and Machine Learning (ML)

As Kjeldskov et al. 2004 and Joe et al. 2015 before them, Johannsen et al. 2019 find the traditional TA too labor- and time-demanding. Their application area is Continuous Software Engineering (CSE), where developers distribute software updates as soon as they are ready-for-use.

The authors automatize the examination of data with ML methods using own manually transcribed verbal protocols as training data. They also set up a classifier prototype of their own. It classifies according to sentiments: positive, negative, neutral, unclear. User behavior is checked for these features.

The classifier is delivering results in short time. Where users react in unclear or negative fashion, the developers should improve their work.

Goldberg et al. 2020 experiment with machine learning, too. They check audio takes from patients and therapists in order to learn something about the alliance of patient and therapist against the disorder. In their training data and learning data they calculate unigrams / bigrams with Sent2Vec and tf-idf .

Fan et al. 2019, 2020a also want to automatize the expensive examination of verbal protocols using ML methods. They start out with 12 transcribed verbal protocols with additional features like sentiment or pitch. The recordings are submitted to raters who tag the interesting protocol sequences with the categories reading, procedure, observation and explanation. A problem span is defined as a sequence of these categories. Now the mapping of categories and features to transcript spans is calculated according to precision, recall and F-value. As main result the authors can show that considering the combination of all features performs better than relying on the transcript alone.

Thinkie 3

The beginnings

Thinkie supports the data capture and initial analysis in TA studies on iPhone and iPad.

Background is the own experience of the developer. During her field studies on text summarization. I moved around to colleagues in Germany and in the US for capturing their verbal protocols. My equipment included a laptop, a mobile printer, an audio recorder, and a lot of paper - result in Endres-Niggemeyer 1998. A handy mobile technical outfit performing everywhere would have been better, light and small enough to fit into a lady's handbag. At that time, nothing like this was available. Well, if it does not exist, develop it yourself. This can be done on an iPhone and an iPad.

A first Thinkie version appeared in 2013 (Endres-Niggemeyer 2013). It presented the approach.

Thinkie 2 came out somehow in 2015. It advanced towards real usability and adapted to the current IOS system version. Speech recognition improved. A CloudDrive access was provided.

Thinkie 3 adapts to the current IOS of 2021. Speech-to-Text uses the up-to-date IOS Speech API. All data is stored in the user's private iCloud container. From there it can be exported. Thinkie 3.1 includes an explicit user authorization with the AppleID.

In the following Thinkie 3 will be described. Its main features:

- Mobility: handbag size on iPhone and iPad
- Working worldwide in IOS
- iCloud connection
- Data storage in the user's private iCloud-Container
- Own data model of cloud files
- Audio recording on the iPhone, video recording on the iPad in parallel
- Analysis / transcription of verbal protocols on the iPad
- Speech Recognition (Speech-To-Text)
- Export of the iCloud files
- Colorful user interface

Thinkie 3 on iPhone and iPad with cloud connection

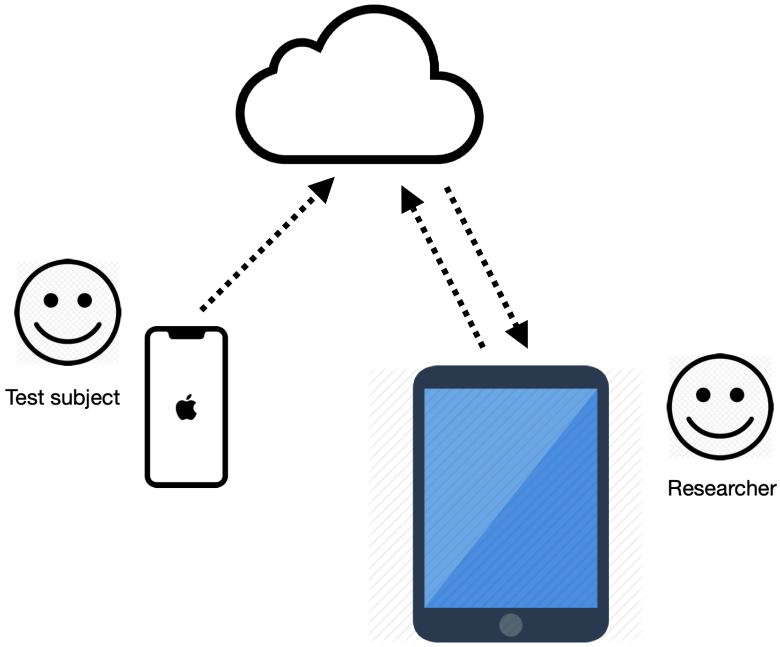

Fig. 7. Mobile working environment: iPhone for the test subject, iPad for the tutor / researcher, cloud connection for the data

For using Thinkie one just needs an iPhone, an iPad and the iCloud connection for both. That's it (fig. 7). Thinkie works where ever a sufficient Internet access is available. The mobility is in particular useful for field research. It is no drawback elsewhere.

With two persons the staff is complete: The test subject uses the iPhone and records the verbal protocol. The researcher disposes the iPad to take the test subject with the surroundings and makes the video recording run.

As soon as the recording is done one may check it and sends both audio and video file one after the other to the cloud container. The researcher will manage the further handling the recorded files on the iPad.

On the iPad one can inspect the verbal protocol and the video, so that test subject and researcher can discuss it immediately. One can transcribe the audio verbal protocol and elaborate it in a simple fashion.

For a more sophisticated interpretation and other purposes all the data files on the cloud can be exported.

Authorization with Apple ID

An explicit authorization with the user’s AppleID is needed in order to make the connection to the private iCloud container of the user work.

Thinkie asks for the AppleID right at the start. The authorization page conforms to local conditions. The fig. 8 shows a version using the device code of the iPad. A face recognition option might run as well.

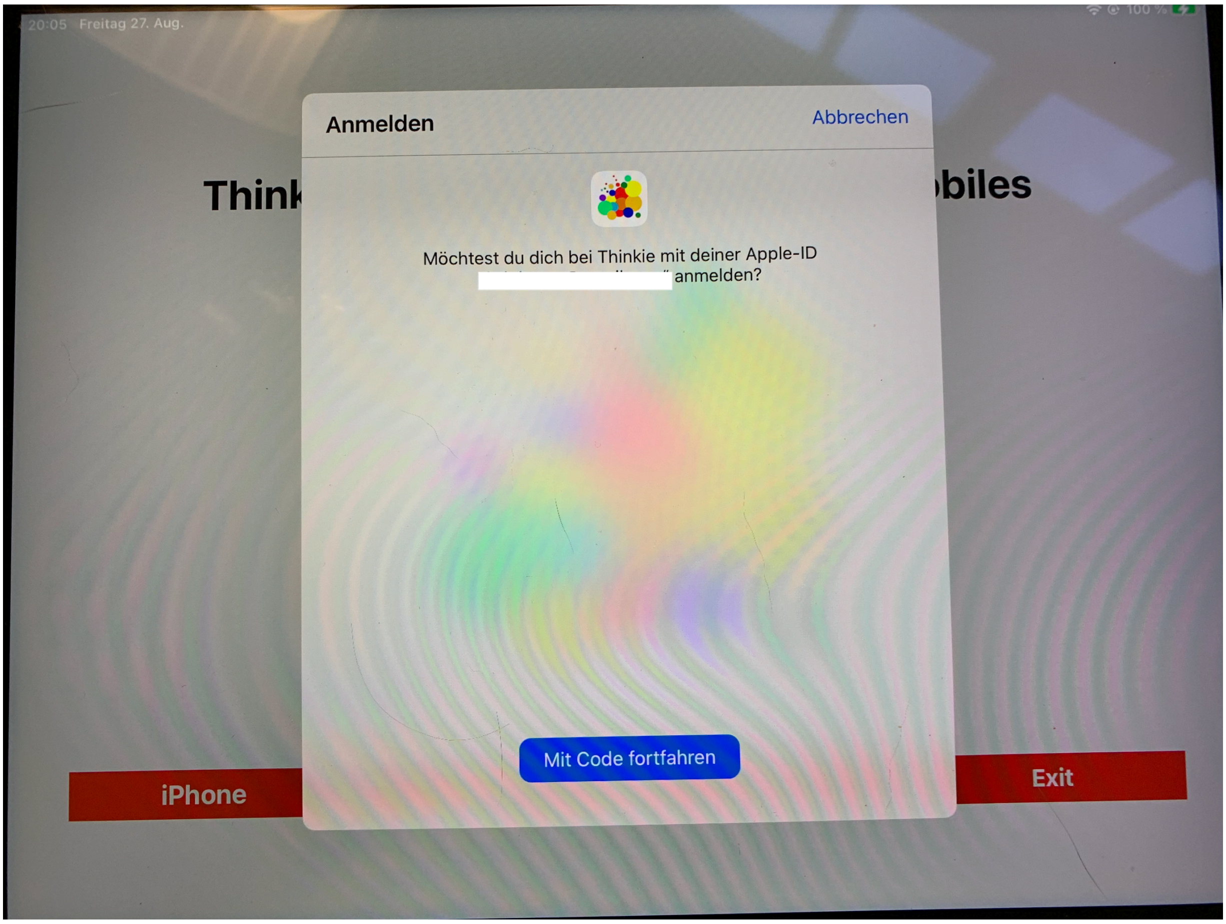

Abb. 8. Localised AppleID authorization with device code on iPad

ThinkTon: Thinkie's data model in the iCloud container

ThinkTon

System fields

recordName	Reference
createdAt	Date/Time
......	

Custom fields

recordID	String
kennung	String
mode	String
study	String
subject	String
title	String
audioAsset	Asset
audioDuration	Double
transcriptAsset	Asset
videoAsset	Asset
startClips	Double(List)
endClips	Double(List)

Fig. 9. ThinkTon - the data model of Thinkie 3 in the user's private iCloud container

A data model called ThinkTon (fig. 9) in the user's private iCloud container stores Thinkie's permanent data. The iPhone writes ThinkTon records into the sector "phone" of the container, the iPad uses the sector "pad".

From the system data fields, only the identifier record name and the creation date „createdAt" are used.

In the custom area „study", „subject" and „title" store three parts of the record name. The three strings are concatenated to get the "kennung" (= identifier) property, the name of the record. It is used for record retrieval.

The data files of the investigation are stored as assets. Assets accept large data sizes. There is one for audio data, one for video data and one for transcript data.

The duration of the audio file is stored separately, the video duration is taken from the video file itself.

The “startclips” and “endclips” lists contain the starts and ends of audio clips if the audio file was segmented into manageable spans for manual transcription (see fig. 20).

A clip is a timestamp of the data type Double (“double-precision, floating-point value“).

Thinkie starting up from iPhone and iPad

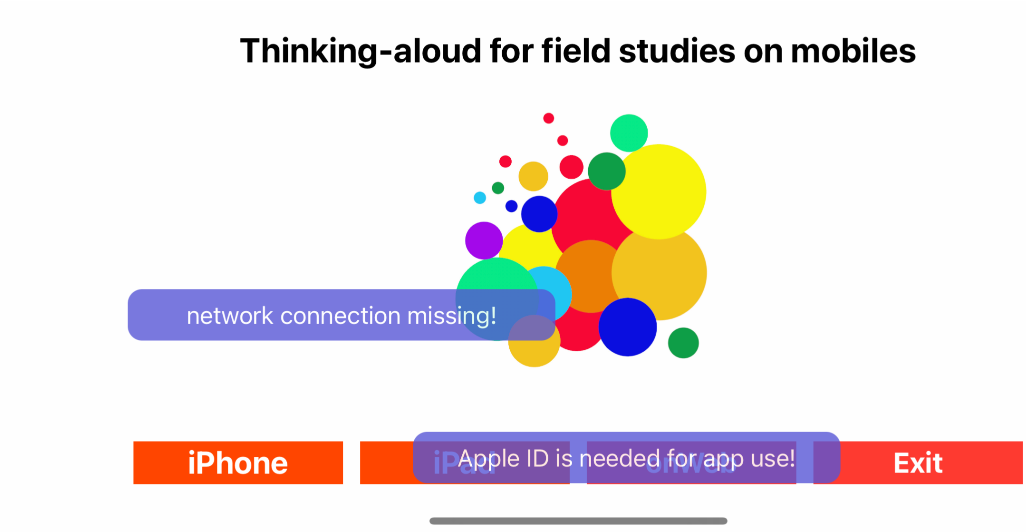

Fig. 10. Shared Thinkie start screen on iPhone and iPad: currently no network and no Apple ID authorization

Thinkie 3 starts up with a screen that iPhone and iPad share (fig. 10). If the operation conditions of the app are not set (as in fig. 10) blue error messages pop up as shown in the figure. The links to the iPhone and the iPad are blocked. Also elsewhere in the app blue notes inform the user of some need for action.

The “onWeb” button leads to background information, Including user instructions in English and German.

“Exit” switches back to IOS.

On the iPhone

On the iPhone, Thinkie offers one user interface only. Its functions are shown in fig. 11, the user view in fig. 12.

The iPhone serves for recording what the test person says. There is no restriction of languages. The resulting sound file contains the spoken verbal protocol. Noises are recorded as well.

The sound file needs a name of three items called “study”, “title” and “subject”. Users fill in any strings. The items are concatenated to get the record identifier. It is used for retrieving the record.

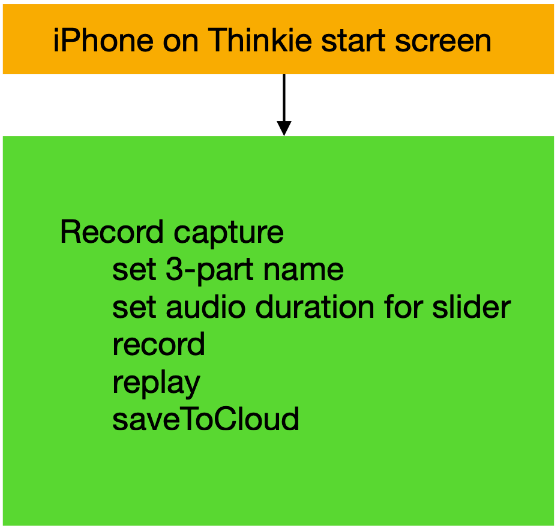

Fig. 11. Functional view of the iPhone user interface

A slider shows how the capture advances (fig. 12). Its expected maximal duration can be set. If the recording takes more time, the slider stops, but the recording goes on.

When the recording is stopped, one can listen to it. If it is flawed, one may prefer to replace it. Otherwise one saves the audio file into the cloud container. A “stored!“ message indicates that it arrived ok, otherwise an error message appears.

For the iPhone owner the recording permission is preset. “Recording allowed” confirms this. Only at the very first call the user is asked to permit the use of the microphone (fig. 13). Without consent no capture.

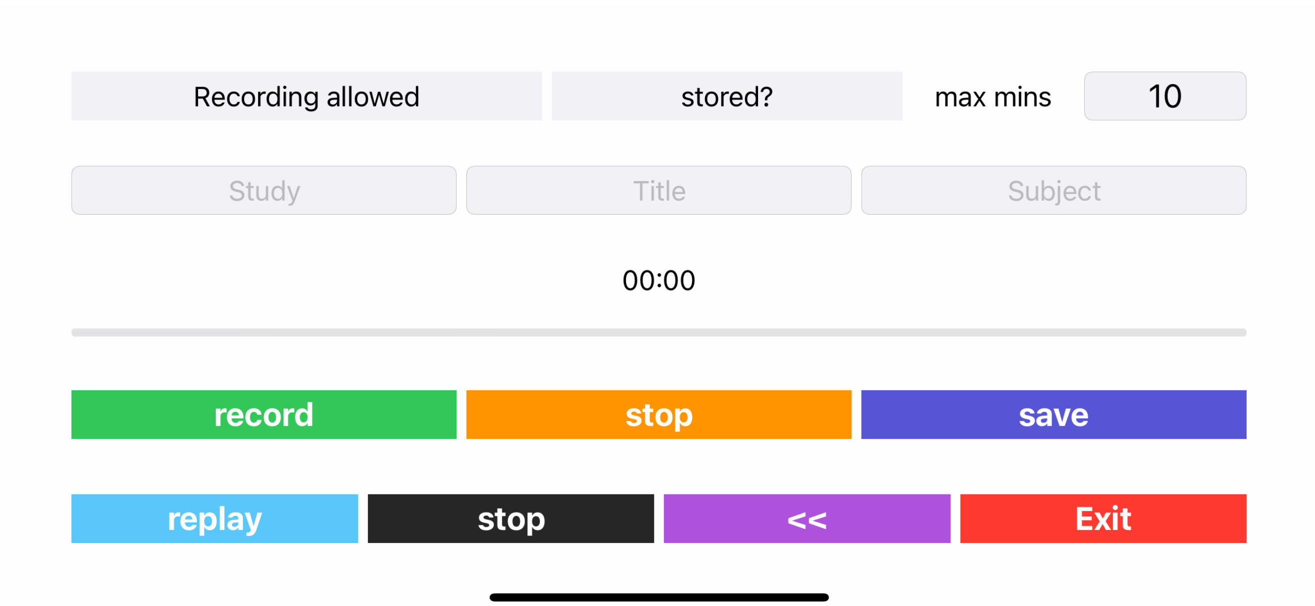

Fig. 12. iPhone user interface in start state

Fig 13. Asking for microphone use permission (localized)

On the iPad

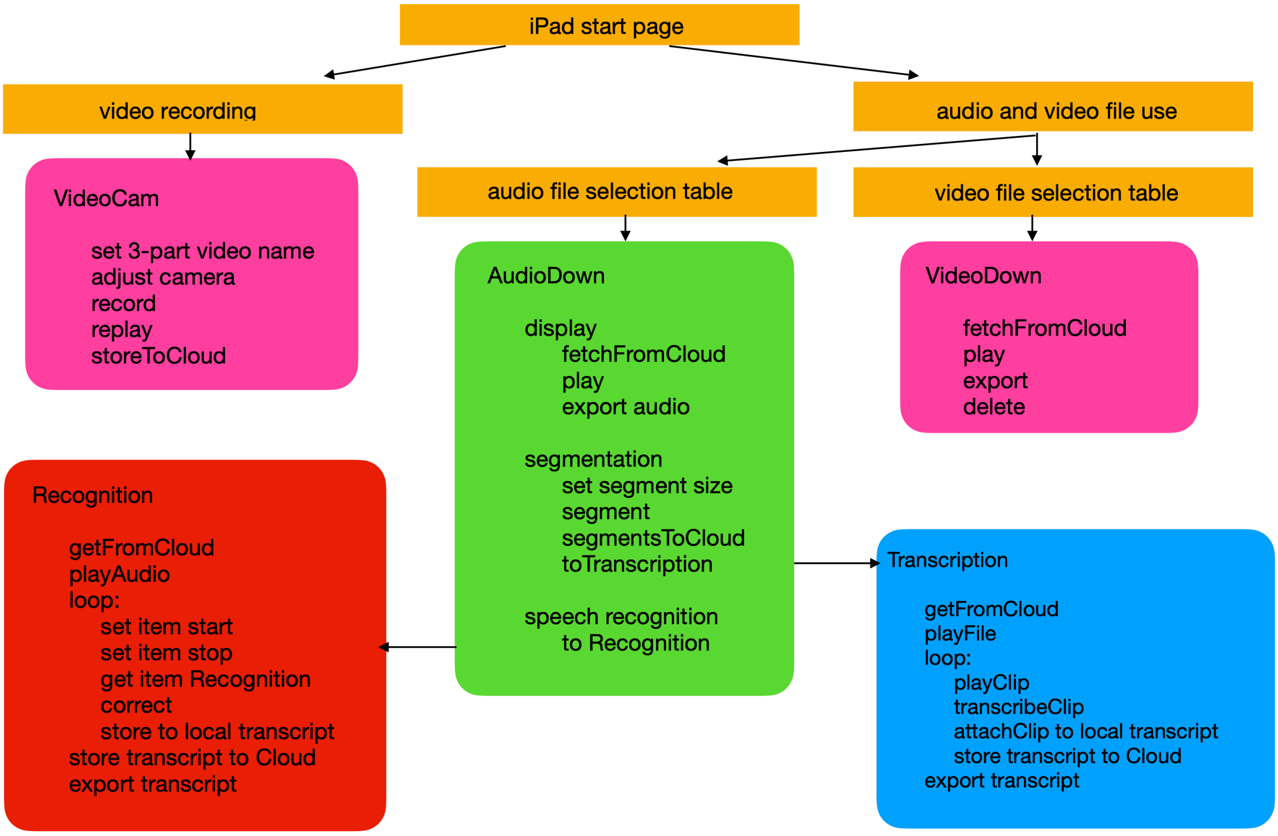

Fig. 14. Functional display of Thinkie on the iPad

Seen from outside (fig. 7) the iPhone and the iPad may appear equal. However the functional view (figs. 11 and 14) shows that they are not: the iPad has a much higher functional load than the iPhone.

On the iPad, the tasks are distributed on several user interfaces which are reached through intermediate screens (in yellow in fig. 14). They communicate with the cloud container to store and fetch their data.

Videos are captured on the video page “VideoCam” (fig. 17). They are stored in the cloud container. For playing and exporting a video, one fetches it on the page “VideoDown” (fig. 18).

The audio processing page “AudioDown” (fig. 19) leads to two subviews for manual transcription (“Transcription” - fig. 20) and to the page for Speech-To-Text (“Recognition” - fig. 21).

Because of different demands the audio processing follows two functional lines:
- In case of segmentation for manual transcription the start and end times of the segments are stored. The “Transcription” page offers the segments on a table so that they can be handled one by one. One writes the text into the transcript.
- On the “Recognition” page the user selects sound intervals that make sense and sends them to the Speech recognition server. Its output is attached to the transcript.
- The Transcription and the Recognition page share the transcript file store in the cloud container. They store their data there and fetch it again from there.

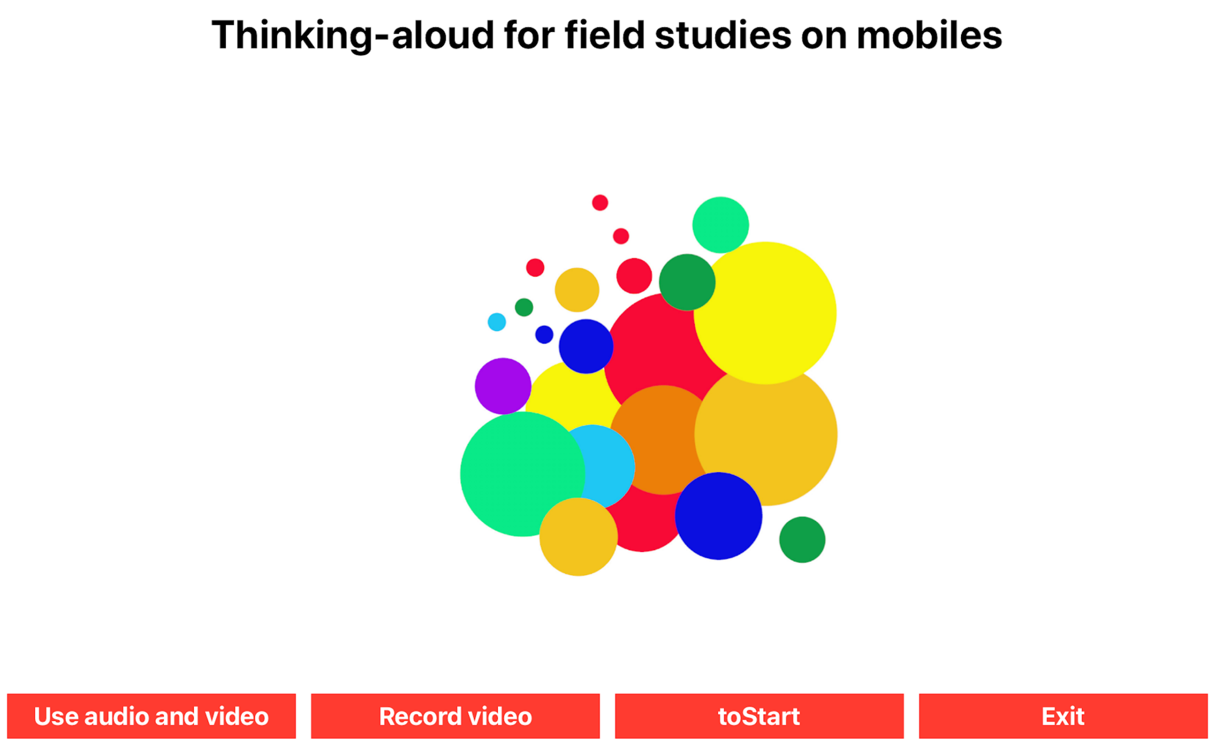

Fig. 15. Initial user interface on the iPad

On the iPad start page (fig. 15) the button “Record video” leads to the interface for video capture. It functions in the same style as the audio recording described above. The video needs a three-items name. The forefront and the back-side camera can be used. After video take one can replay it and store it in the cloud container. More detail is explained in the user guides (in English and German).

The button „Use audio and video“ switches to a table listing the available audio and video files (fig. 16). Each entry of the table presents a button that fetches the file.

The button “toStart” switches to the common first interface of the iPhone and iPad, “Exit” stops the app and returns to IOS.

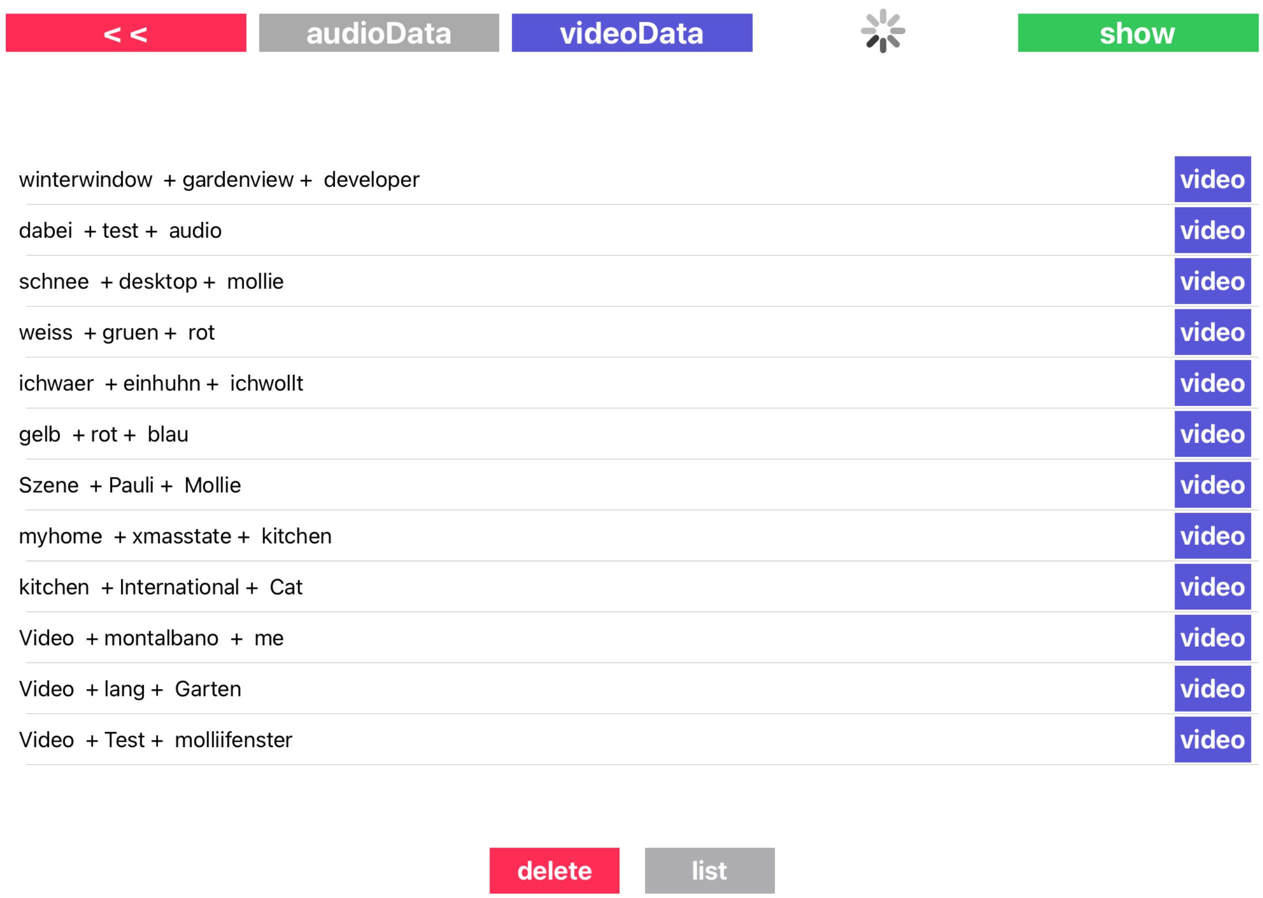

Fig. 16. Selection table of video files with alternative switch for audios

Video recording and playing

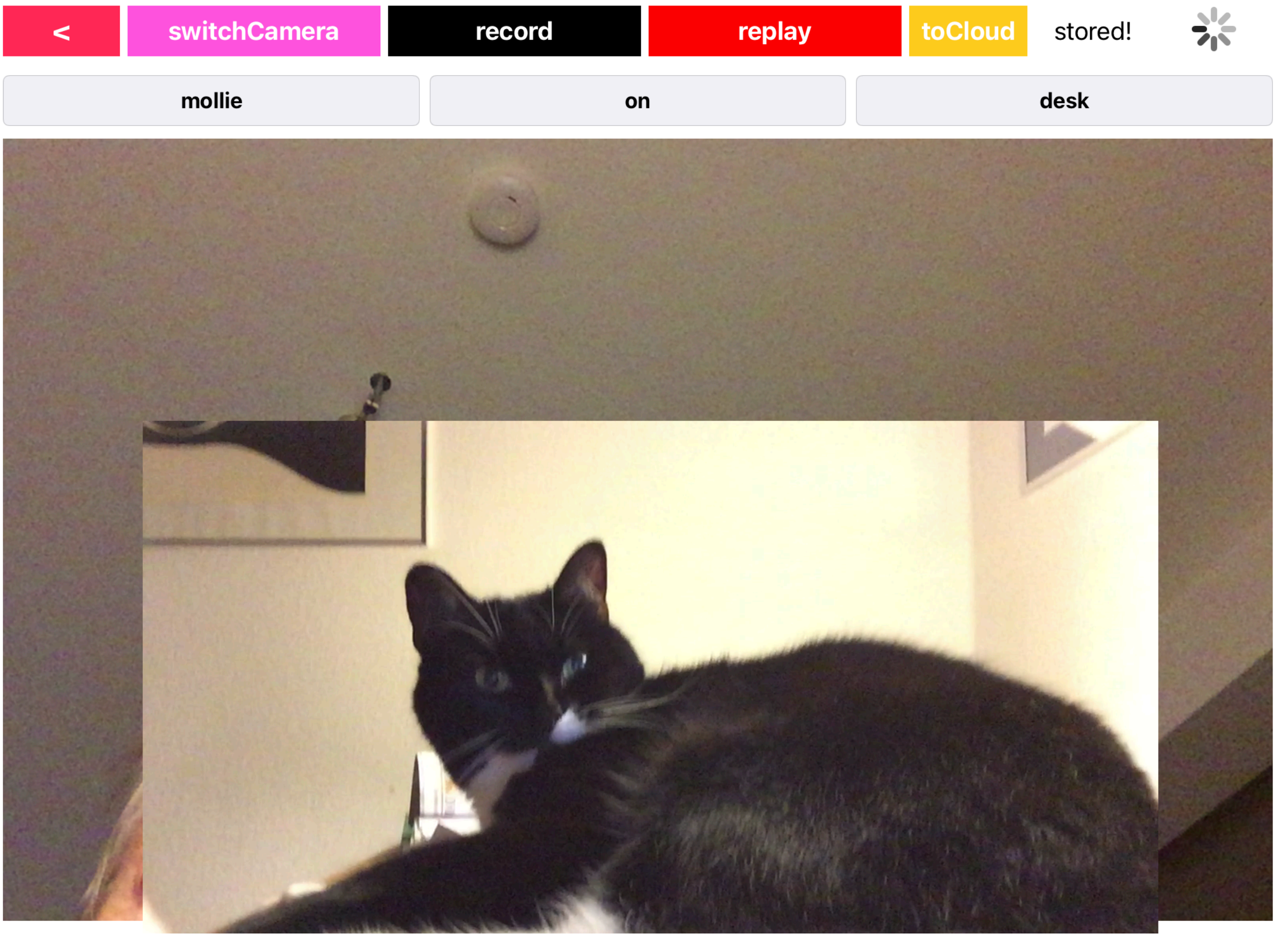

Fig. 17. Screen for Video recording: buttons, three-part name, current camera view with replaying video in front

In parallel with recording audio on the iPhone one can take a video on the iPad. The screen for video recording is reached with the “record Video” button on the iPad start page.

Fig. 17 shows what works for video takes.

The camera can be switched to the front or to the rear side of the iPad. The “record” button starts and stops the recording.

The video file got a three-part name.

On fig. 17 the video has already been recorded. The file has been saved to the cloud container. This is confirmed by the “stored!” entry.

Currently the video is replaying in the foreground. It is available as long as it is not overwritten.

Behind one sees the current preview of the camera. From there a new take might start.

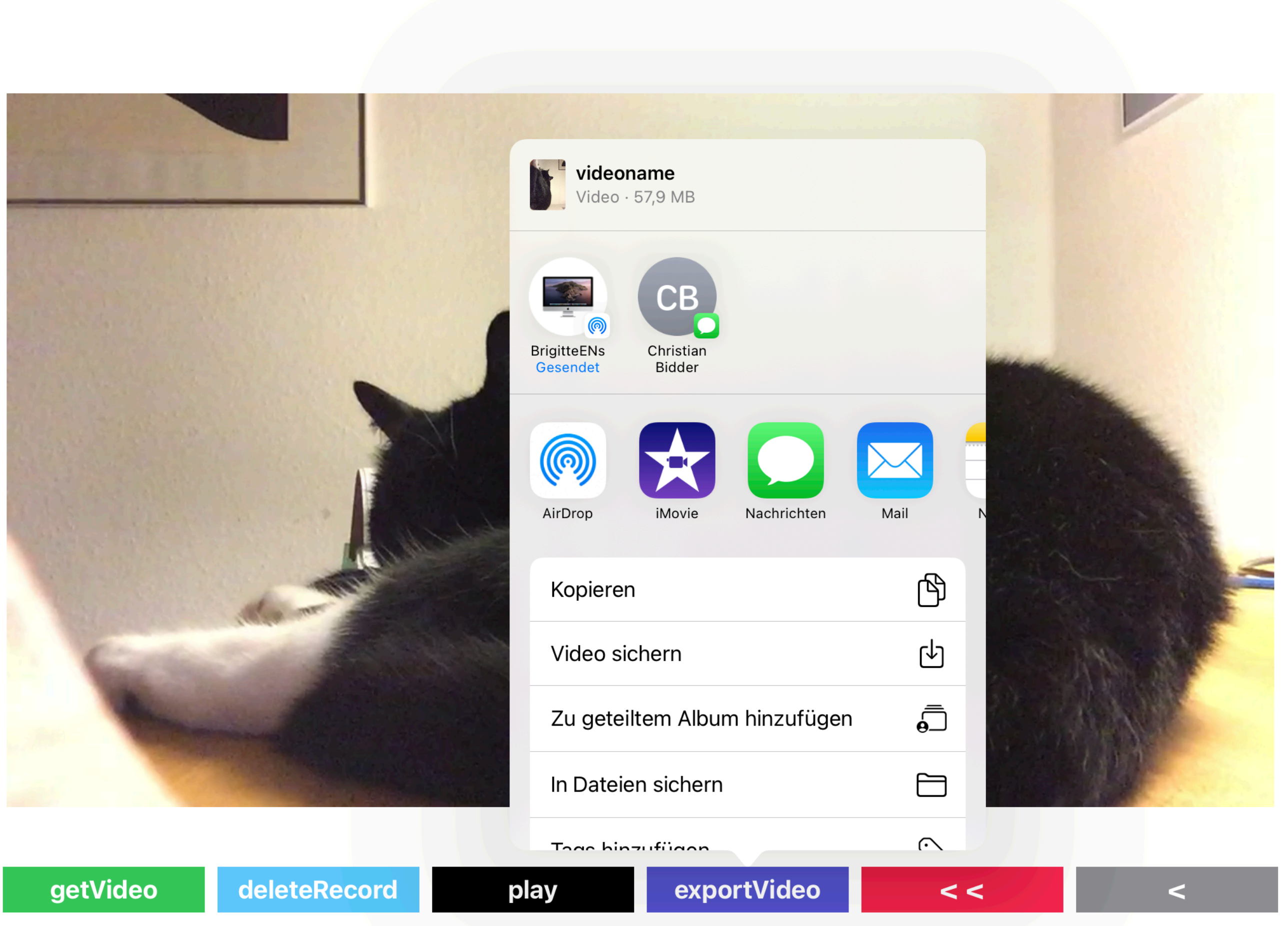

Fig. 18. VideoPlay page: download, play, export and delete a video

The videos on the cloud container are presented when one selects the "videoData" switch on the overview list (see fig. 16). The buttons on the right lead to the chosen video.

Fig. 18 illustrates what can be done with the selected video file (cf. fig. 17).

First one downloads the video file from the cloud container ("getVideo"). As soon as it has arrived the buttons for its treatment activate. One may play the movie ("play"), export it ("exportVideo"), and delete it ("deleteRecord").

When the fig. 19 was shot the video is being exported into the system environment of the user. Indeed it has already arrive on the developer's iMac.

Using "deleteRecord" cancels the video file in the cloud container. A message pops up when the delete succeeded.

The buttons on the right below return to the video/audio overview table and to the iPad start.

Speech processing : overview

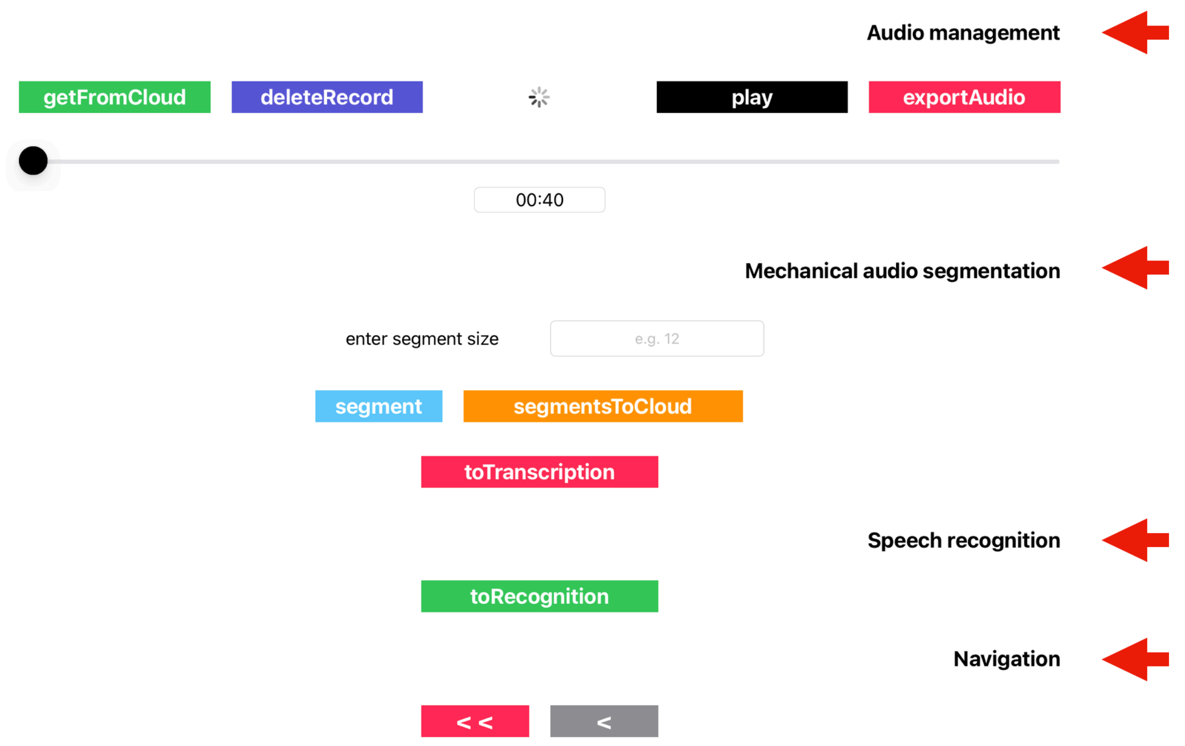

Fig. 19. Speech-To-Text: Distribution page of speech processing

The Apple Speech API for speech recognition handles the basic transcription of the sound files. It integrates many resources of Natural Language Processing. Thinkie 3 uses its server-based version because it processes more languages and achieves better results than the alternative solution on the device itself.

Even if one restricts the demands on Speech-To-Text - the transfer from spoken to written text - this will often turn out to be problematic. This is why Thinkie offers a mixed approach of automatic support and human intellectual transcribing and control. All transcription results are collected in the transcript file.

The NLP distribution page (fig. 19) supports straightforward services:
- handling the audio file: download it, play it, export it and delete it.
- cutting manageable clips from the audio file. The audio clips of a chosen duration can be saved in the cloud container. They are presented on the transcription page (fig. 20).
- navigating back to the list of files and to the iPad start page.

Speech recognition handling is more complicated for the user. Therefore it has a page of its own (fig. 21).

Speech-To-Text often means managing uncertainty. Who knows the task cannot ignore this. Some more detail is described above in the transcription section.

The Speech API adapts to the speaker, but often it has to interpret what a given sound sequence might mean in written text. The interpretation is easier when the input units are interpretable and make sense.

Thinkie asks users to pick meaningful passages with the tools on the recognition page.

Segmentation

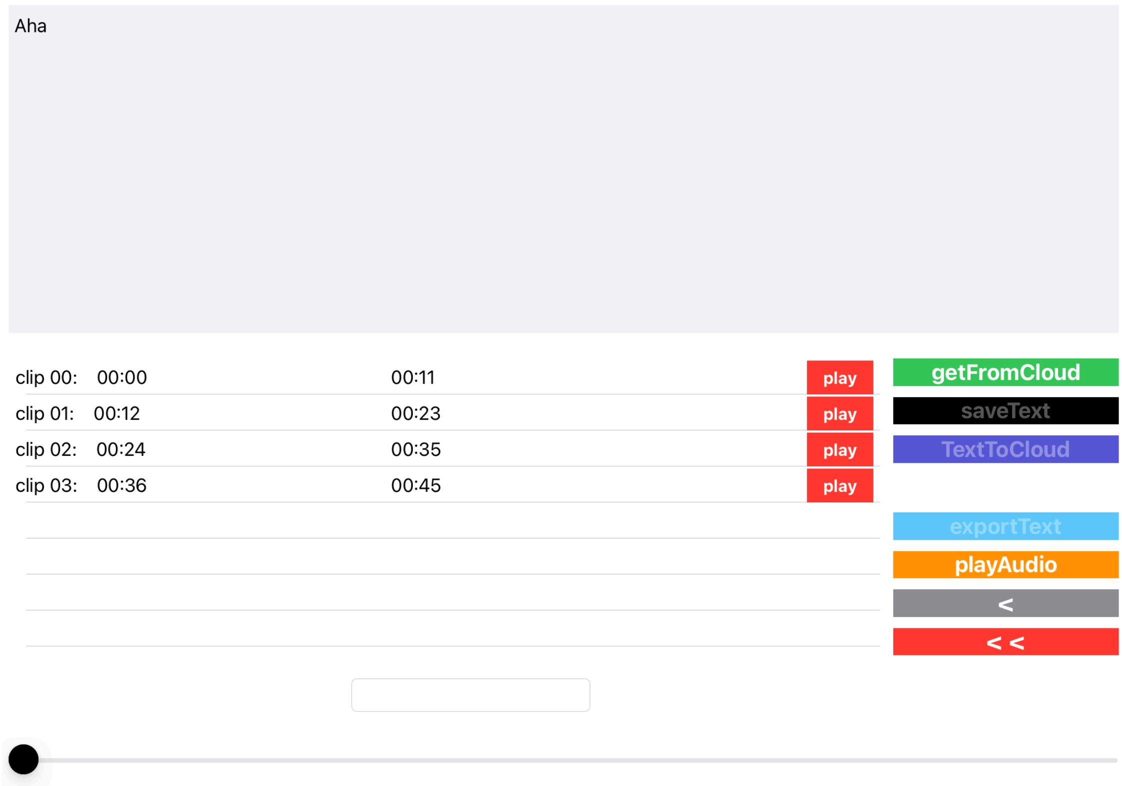

Fig. 20. Segmentation page: target transcript, list of audio clips, buttons for general tasks, time label and slider

The segmentation page represented on fig. 20 shows at its top the transcript area (still empty).

Using the “getFromCloud” button one inserts the transcript from the cloud container into the transcript area and the audio clips into the table below.

Now one can play the audio clips one by one using the “play” button at their right. The reached position shows up in the time label below. One can add and change passages of the transcript - correction of recognition results included.

The transcript can be stored locally (“saveText”) and in the cloud container (“TextToCloud”).

The “playAudio” button plays the whole file. It is accompanied by the slider at the page bottom.

The last buttons return to the audio/video file list and to the iPad start screen.

Speech recognition - Speech-To-Text

For Speech-To-Text - speech recognition - one needs meaningful takes from the audio file that do not overcharge the recognition engine. Their duration should not exceed 60 seconds. Takes that make sense are recognized better.

The recognition engine is set to the system language of the device. Changing it eg. from German to French will make the recognition switch from German to French. The user has to agree to speech recognition explicitly.

Thinkie picks from the sound file a clip specified by the user and sends it to the recognition server. As soon as the result comes in it is attached to the transcript. One can check and rework it there.

The clips are picked one after the other from the playing audio file by managing the “play” and “pause” buttons (cf. fig 21).

The current start and stop values are displayed on the interface - the start field shows the start time of the clip, the stop field its stop time. With the next clip of the sound file, the old stop value moves to the start field, and the new stop value is entered.

With the “recognize” button the audio interval from start to stop is sent to the recognition server. As soon as the recognition result has been attached to the transcript one can pass to the next audio clip.

Practically speaking, for a recognition action:

1. you press „play“
2. you press „stop“ - the stop value appears in its field

Then for every recognition unit:

3. you press „play“ again - the old stop value moves to the start field
4. you press „stop“ - a new stop value appears in the stop field
5. you click „recognize“ for the sequence from start to stop
6. you wait for the recognition result to be appended to the transcript, you may correct it if necessary
7. back to step 3. - handle the next sequence

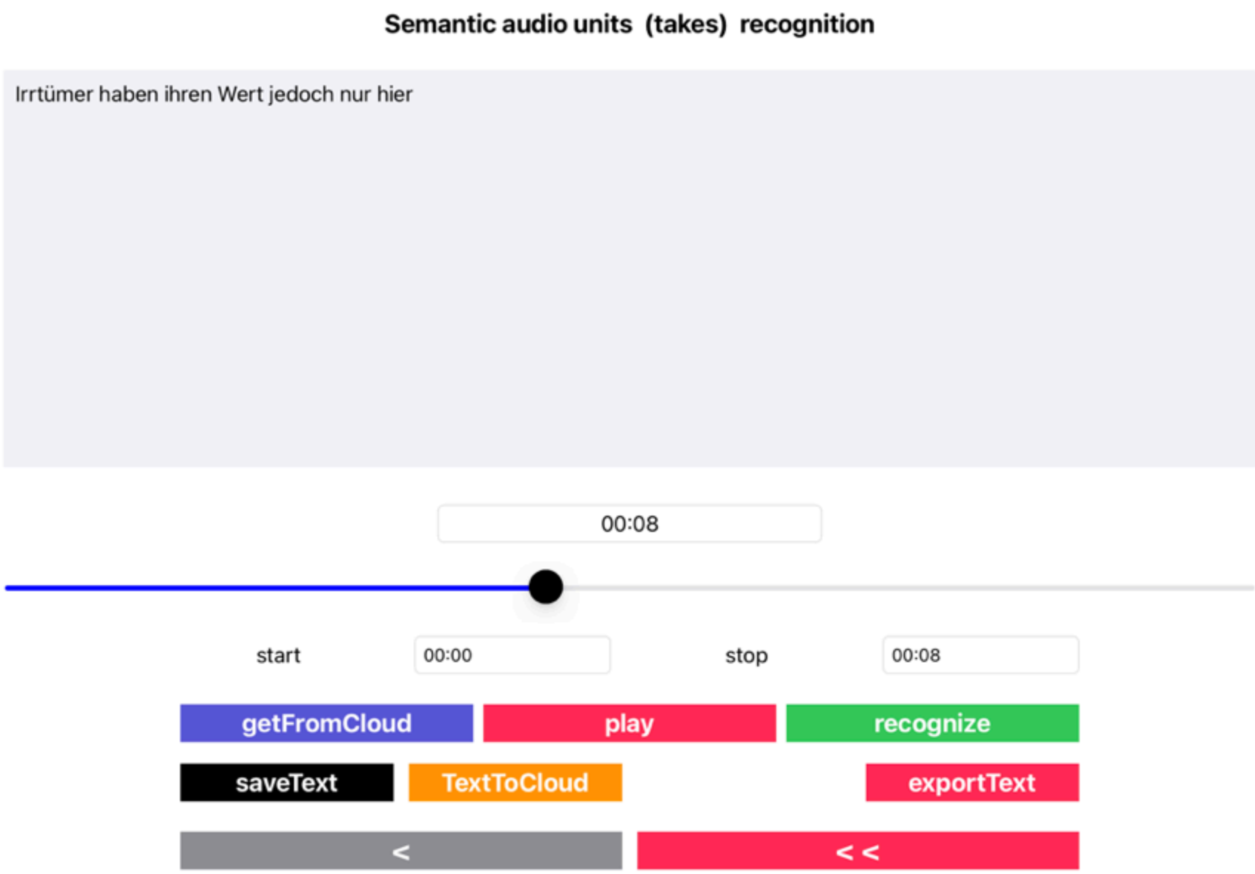

Fig. 21. Speech-To-Text is under way

On fig. 21 a situation during the speech recognition run is presented.

The user has managed several steps before:

- The audio file was downloaded from the cloud container („getFromCloud“)
- The audio file was started („play“) and stopped immediately, so so that the start time for recognition was set to „00:00“.
- The user played on until a meaningful sequence was reached and stopped. The stop time for the recognition is set to „08:00“.
- By clicking the „recognize“ button the user sent the specified sequence to the speech recognition server.
- The server answered. The text corresponding to the sound has been entered into the transcript field.

For obtaining the next stretch of text, the user presses “play” (step 3 of the loop) and listens for up to 60 seconds till the next meaningful passage has been played. Then she stops and sends the current clip to the recognition - and so on till the end of the file is reached.

Conclusion

Thinking Aloud

Thinking Aloud of 2021 has grown far out of what the founding fathers Ericsson & Simon conceived it to be in 1980.

TA has spread to many branches. Some of them did not yet exist in 1980. TA procedures adapted to the local needs.

The Internet of today offers possibilities that one could not imagine in the early stages of TA.

Nowadays audio and video data capture is the standard, possibly combined with eye tracking. One can compare what is said and what is done. More data of better quality is available, engendering a better basis for data analysis.

The thinking acts detected in a verbal protocol have been equipped with an internal structure, a data model of their own. Data models vary. They contain properties corresponding to the aims of a specific research.

Automatic Speech-To-Text (speech recognition and text analysis) can participate in the labor-intensive transcription and mitigate the work load. Transcription can be outsourced to external service providers.

Methods for coding and interpretation of verbal protocols are available.

There are first applications of ML (Machine Learning) classifiers.

The user interfaces and their users have multiplied since 1980. How users manage the interfaces is much more an issue than in former times. It is handled in User Experience (UX) research. Like this TA for learning how users get along has become more significant in Usability Engineering.

What Thinkie 3 offers

Thinkie 3 offers an integrated handbag-size mobile support for Thinking Aloud. It is useful for field research, but it works just wherever a sufficient network connection is available.

Thinkie is running world-wide in IOS on cooperating iPhone and iPad pairs. The data is stored in a cloud data structure.
Thinkie 3 concentrates on data capture with audio and video and a simple transcription of verbal protocols with the help of the server-based Apple Speech API.

All data can be exported as required for a more thorough data exploitation.

So far there are no comparable apps for TA on the IOS app store.

Going on?

For the developer it is a practical issue whether she should envision a next Thinkie version.

In the scientific and technical environment around no stagnation is observed, so one might contribute to it . Natural Language Processing (NLP) and its applications advance. The user perspective gains more interest. This affects TA positively.

Improvements for Thinkie 4 are easy to imagine: extending the Speech API use, set up a general scheme for semantic data structures, support more data analysis than simple transcription, and so on.

Beyond this, the target is keeping track with IOS development and using new technical possibilities.

If others had surpassed Thinkie one might leave the topic all together to others. But nothing like this on IOS.

So let us see!

References

Aitken LM, Mardegan KJ (2000). “Thinking Aloud”: Data Collection in the Natural Setting. Western Journal of Nursing Research, 2000, 22(7), 841-853 .

Al-Janabi H, Keeley T, Mitchell P, Coast J (2013). Can capabilities be self-reported? A think aloud study. Soc Sci Med. 2013 Jun; 87: 116–122. doi: 10.1016/j.socscimed.2013.03.035 .

Alhadreti O, Mayhew P (2018). Rethinking Thinking Aloud: A Comparison of Three Think-Aloud Protocols. 2018 CHI Conference on Human Factors in Computing Systems. Computers in Human Behaviour 19 (2018) 95 -115. https://ueaeprints.uea.ac.uk/id/eprint/65771/1/paper1268.pdf

Amin HU, Malik A (2013). Human memory retention and recall processes. Neurosciences 2013; Vol. 18 (4): 330-344. https://nsj.org.sa/content/nsj/18/4/330.full.pdf

Asselin M, Moayeri M (2010). New tools for new literacies research: an exploration of usability testing software. International Journal of Research & Method in Education 33(2010)1, 41–53.

Baddeley A (2003). Working memory: looking back and looking forward. Nat Rev Neurosci 4: 829-839. https://studentski.net/gradivo/ulj_fif_fi2_fkz_sno_delovni_spomin_01

Bernardini S (2001). Think-aloud protocols in translation research. Achievements, limits, future prospects. *Target* 13:2 (2001), 241–263. https://www.researchgate.net/publication/233632990

Boren MT, Ramey J (2000). Thinking Aloud: Reconciling Theory and Practice. IEEE Transactions on Professional Communication 443(3) 261-278. https://www.researchgate.net/publication/3230127

Cowan N (2010). The magical mystery four: How is working memory capacity limited, and why? *Curr. Dir. Psychol. Sci.* 2010;19(1):51–57.

Crutcher RJ (2007). CAPAS 2.0: A computer tool for coding transcribed and digitally recorded verbal reports. Behavior Research Methods 39(2007)2,167-174

Dahl AC, Carlson SE, Renken M, McCarthy KS, Reynolds E (2021). Materials Matter: An Exploration of Text Complexity and Its Effects on Middle School Readers' Comprehension Processing. LSHSS 52(2021)2, 702-716, https://pubs.asha.org/doi/10.1044/2021_LSHSS-20-00117.

Dresing T, Pehl T, Schmieder, C (2015): Manual (on) Transcription. Transcription Conventions, Software Guides and Practical Hints for Qualitative Researchers. 3rd English Edition. Marburg Available Online: http://www.audiotranskription.de/eng-lish/transcription-practicalguide.htm .

Elzakker van CPJ M, Delikostidis I, van Oosterom PJM (2008). Field-Based Usability Evaluation Methodology for Mobile Geo-Applications. The Cartographic Journal 45(2008)2,139–149. https://www.researchgate.net/profile/Peter-Oosterom/publication/233646203

Endres-Niggemeyer B (1998): Summarizing Information. Heidelberg: Springer.

Endres-Niggemeyer B (2000). SimSum: an empirically founded simulation of summarizing. Information Processing] Management 36(4)659-682 .

Ericsson KA (2006). Protocol Analysis and Expert Thought: Concurrent Verbalizations of Thinking during Experts’ Performance on Representative Tasks. in: The Cambridge Handbook of Expertise and Expert Performance, 223-242.Cambridge: Cambridge University Press. doi:10.1017/CBO9780511816796.013

Ericsson KA, Simon, HA (1980). Verbal reports as data. Psychological Review, 87(1980)3, 215–251. https://doi.org/10.1037/0033-295X.87.3.215

Fan M, Li Y, Truong KN (2020a). Automatic Detection of Usability Problem Encounters in Think-aloud Sessions. ACM Trans. Interact. Intell. Syst. 10(2)Article 16. https://dl.acm.org/doi/10.1145/3385732 .
Fan M, Lin J, Chung C, Trong KN (2019). Concurrent Thinking-Aloud Verbalizations and Usability Problems. ACM Trans. Compute.-Hum. Interact. 26(5), Article 28. https://dl.acm.org/doi/10.1145/3325281 .
Fan M, Shi S, Truong KH (2020b). Practices and Challenges of Using Think-Aloud Protocols in Industry: An International Survey. Journal of Usability Studies 15(2)85-102 .
Feng Z, González VA, Trotter M, Spearpoint M, Thomas J, Ellis D, Lovreglio R (2020). How people make decisions during earthquakes and post-earthquake evacuation: Using Verbal Protocol Analysis in Immersive Virtual Reality. Safety Science 129(2020), https://doi.org/10.1016/j.ssci.2020.104837, https://www.researchgate.net/publication/341554201
Fischer T, Reinhardt G, Stumpf P, Schwarz PEH, Harst L (2019). Think-aloud usability testing of an app to increase physical activity. https://www.researchsquare.com/article/rs-7670/v1.
Flick U (2019). An Introduction to Qualitative Research. 6th ed. SAGE Publications Ltd.
Freiling I (2019). Detecting misinformation in online social networks: A think-aloud study on user strategies. Studies in Communication and Media 8(4),471-496 . https://www.researchgate.net/publication/338087845
Ghenai A, Smucker MD, Clarke CLA (2019). A Think-Aloud Study about Medical Misinformation in Search Results.ROME 2019, July 25, 2019, Paris, France . https://rome2019.github.io/papers/Ghenai_Smucker_Clarke_ROME2019.pdf
Gibson JJ (1986). The Ecological Approach to Visual Perception. Hillsdale NY, Erlbaum.
Gizzi MC, Rädiker S (2021). The Practice of Qualitative Data Analysis. Research Examples Using MAXQDA. https://www.maxqda-press.com/catalog/books/the-practice-of-qualitative-data-analysis .
Green A (1997). Verbal Protocol Analysis in Language Testing Research. Cambridge, Cambridge University Press. https://www.cambridgeenglish.org/Images/329137-studies-in-language-testing-volume-5.pdf
Han KJ, Kim HS, Kim MJ, Hong KJ, Park S, Yun SN, Song M, Jung Y, Kim H, Kim DOD, Choi H, Kim K (2007). Thinking in Clinical Nursing Practice: A Study of Critical Care Nurses' Thinking Applying the Think-Aloud, Protocol Analysis Method. Asian Nursing Research 2007(1)68-82 . https://www.researchgate.net/publication/248605112
Hughes J, Parkes S (2003). Trends in the use of verbal protocol analysis in software engineering research. Behavior & Information Technology 22(2003)2,127-140. https://www.researchgate.net/publication/220208606
Hutchison P, Monaghan I, Morgan R (2015). A multidimensional analysis method for think-aloud protocol data . PERC 2015 Proccedings. https://www.researchgate.net/publication/301451748
Jackson K, Bazeley P (2019). Qualitative Data Analysis with NVIVO. 3rd. ed., Sage Publications Ltd.
Jaspers M (2009). A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence. International Journal of Medical Informatics 78(2009)340-353.

Jaspers MWM, Steen T, vandenBos C, Geenen M (2004). The think aloud method: a guide to user interface design. International Journal of Medical Informatics (2004) **73**, 781—795 .

Joe J, Chaudhuri S, Le T, Thompson H, Demiris G (2015). The use of think-aloud and instant data analysis in evaluation research: Exemplar and lessons learned. Journal of Biomedical Informatics 56 (2015) 284–291 . https://www.researchgate.net/publication/278161764

Johanssen JO, Reimer LM, Bruegge B (2019). Continuous Thinking Aloud. 2019 IEEE/ACM Joint 4th International Workshop on Rapid Continuous Software Engineering and 1st International Workshop on Data-Driven Decisions, Experimentation and Evolution (RCoSE/DDrEE). https://ase.in.tum.de/lehrstuhl_1/research/paper/johanssen2019continuousThinkingAloud.pdf

Kellogg RT (1987). Effects of topic knowledge on the allocation of processing time and cognitive effort to writing processes. Memory & Cognition 15(1987)3, 256-266 . https://link.springer.com/content/pdf/10.3758/BF03197724.pdf

Kjeldskov J, Skov M B (2014). Was it Worth the Hassle? Ten Years of Mobile HCI Research Discussions on Lab and Field Evaluations. 16th International Conference on Human-Computer Interaction with Mobile Devices and Service, Mobile HCI 2014 (pp. 43-52). https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.700.6445

Kjeldskov J, Skov M B, Stage J. (2004). Instant Data Analysis: Evaluating Usability in a Day. In *Proceedings of NordiCHI 2004* (pp. 233-240). Association for Computing Machinery. https://doi.org/10.1145/1028014.1028050

Klein GB (2015). Verbal interactions and their transcription. https://lnx.keyandkey.it/wp-content/uploads/2020/04/KLEIN-Guidelines-for-transcription-RADAR.pdf

Lester NL, Choo J, Lochmiller CR (2020). Learning to Do Qualitative Data Analysis: A Starting Point. Human Resource Development Review 19(2020)1,94 - 106.

Levy CM, Ransdell S (1994). Computer-aided analysis of writing processes. Behavior Research Methods, Instruments, & Computers 26(1994)2,219-223 . https://link.springer.com/content/pdf/10.3758/BF03204625.pdf

MacLin OH, MacLin MK (2005). Coding observational data: A software solution. Behavior Research Methods 37(2005)2, 224-231 . https://www.researchgate.net/publication/7592953

Maher C , Hadfield M, Hutchings M, de Eyto A (2018). Ensuring Rigor in Qualitative Data Analysis: A Design Research Approach to Coding Combining NVivo With Traditional Material Methods. International Journal of Qualitative Methods 17, 1 - 13.

McDonald S, Edwards HM, Zhao T (2012). Exploring Think-Alouds in Usability Testing: An International Survey. IEEE Transactions On Professional Communication 55(1), 2 - 19 .

Monsalve B, Aliane N, Puertas E, Fernandez Andres J (2020). Think Aloud Protocol Applied in Naturalistic Driving for Driving Rules Generation. https://www.mdpi.com/1424-8220/20/23/6907

Nielsen J (1994) Usability Engineering. San Diego CA, Morgan Kaufmann

NN (2014?). Qualitative Data Analysis Chapter 10. https://www.sagepub.com/sites/default/files/upm-binaries/43454_10.pdf , https://fdocuments.in/document/qualitative-data-analysis-chapter-10.html

Oh K, Almarode JT, Tai RH (2013). An exploration of think-aloud protocols linked with eye-gaze tracking: are they talking about what they are looking at. Procedia Social and Behavioral Sciences 93(2013)184-189 . https://www.researchgate.net/publication/273679982

Olmstedt-Hawala EL, Murphy ED, Hawala S, Ashenfelter KT (2010).Think-Aloud Protocols: A Comparison of Three Think-Aloud Protocols for use in Testing Data-Dissemination Web Sites for Usability. CHI 2010, 2381-2390. https://www.researchgate.net/publication/221518700
Park S, Gooding DC (2014). Working memory impairment as an endophenotypic marker of a schizophrenia diathesis. Schizophr Res Cogn. 2014 Sep; 1(3): 127–136.
Pick I (2015). Das anwaltliche Gespräch. Lang Bern (2015). https://library.oapen.org/handle/20.500.12657/27280
Quak M, London RE, Talsma D (2015). A multisensory perspective of working memory. Front. Hum. Neurosci., 21 April 2015 . https://doi.org/10.3389/fnhum.2015.00197.
Rincon L (2018). Guide for transcribing audio records. https://www.researchgate.net/publication/333059442
Sakar P, Chakrabarti A (2007). A Tool for Supporting Protocol Analysis. ICED'07. https://www.designsociety.org/publication/25325/
Schreiber G, Wielinga B, Breuker J (1993). KADS. A principled approach to knowledge based system development. London: Academic Press.
Selting M, Auer P, Barden B, Bergmann J, Couper-Kuhlen E, Günthner S, Meier C, Quasthoff Uta, Schlobinski P, Uhmann S. (1998). Gesprächsanalytisches Transkriptionssystem (GAT) . https://www.mediensprache.net/de/medienanalyse/transcription/gat/gat.pdf
Shi Q (2010). An Empirical Study of Thinking Aloud Usability Testing from a Cultural Perspective. PhD Thesis Copenhagen Business School. http://hdl.handle.net/10398/8167.
Someren MV van, Barnard YF, Sandberg JAC (1994). The Think Aloud Method. A practical guide to modelling cognitive processes. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.98.7738&rep=rep1&type=pdf, https://www.researchgate.net/publication/215439100
Studer R, Benjamins VR, Fensel D (1998). Knowledge Engineering: Principles and Methods. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.3940&rep=rep1&type=pdf , https://www.researchgate.net/publication/222305044
Sun S (2011). Think-Aloud-Based Translation Process Research: Some Methodological Considerations. *Meta*, *56*(4), 928–951. https://doi.org/10.7202/1011261ar .
Trippas JR, Spina D, Cavedon L, Sanderson M. (2017). A Conversational Search Transcription Protocol and Analysis. In : ACM SIGIR Workshop on Conversational Approaches for Information Retrieval, Tokyo, Japan, August 2017 (CAIR'17). https://www.researchgate.net/profile/Damiano-Spina/publication/319416741
Wang YH (2016). Reading Strategy Use and Comprehension Performance of More Successful and Less Successful Readers: A Think-aloud Study. Educational Sciences: Theory & Practice, 16, 1789–1813.
Yoon YS, Boutin K, Pecaric MR, Fefferman NR, Ericsson KA, Pusic MV (2018). A think-aloud study to inform the design of radiograph interpretation practice. Advances in Health Sciences Education 25(2020)877–903 https://doi.org/10.1007/s10459-020-09963-0 .
Zhou Y, Lin Y (2012). Probe into Translation Process Based on Think-aloud Protocols. Theory and Practice in Language Studies 2(2012)7, 1376-1386. https://www.researchgate.net/publication/260384188