Thinkie 3 für Lautes Denken von 2021

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

BRIGITTE ENDRES-NIGGEMEYER

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

Zusammenfassung

Thinkie 3 unterstützt auf iPhone und iPad das Laute Denken.

Beim Start von Thinkie 3 anno 2021 stellt sich die Frage, in welche Umgebung die App gelangt. Das wird in diesem Artikel zuerst beschrieben. Der aktuelle methodisch-technische Stand des Lauten Denkens von 2021 geht erwartungsgemäß weit über die Anfänge von 1980 hinaus.

Dann folgt eine Beschreibung von Thinkie 3. Als mobiles System speichert Thinkie seine Daten in einem benutzereigenen Cloud-Container. Die Datenstruktur nimmt neben Metadaten Audio- und Videodateien auf, sowie Audioclips und Transkripte. Speech-to-Text von Apple hilft bei der Transkription. Benutzer werden bei der Datenaufnahme und der anfänglichen Datenanalyse unterstützt. Für mehr können die Daten exportiert werden.

Es gibt zusätzlich eine deutsche und eine englische Benutzeranleitung.

Konkurrierende Apps fehlen bis jetzt im IOS. Ein weiter entwickeltes Thinkie 4 ist darum nicht auszuschließen.

Schlagwörter

IOS, iPhone, iPad, iCloud, Lautes Denken, Audio, Video, NLP, Sprache, Datenaufnahme, Transkription, Speech-to-Text, Methoden, Anwendungsgebiete, Benutzeroberfläche, UX, Verbalprotokoll

Überblick

Lautes Denken als Methode gibt es schon länger. Man kann die Grundlagen bei Ericsson & Simon 1980 oder später nachlesen und sich von van Someren et al. 1994 erklären lassen wie man es machen sollte.

Aber auf welchem Stand ist es heute? Die Frage stellte sich, als ich Thinkie 3 für Benutzer in den den AppStore lud, am Ende einer langen Entwicklungsgeschichte. Wie sieht das Umfeld aus, in das die Thinkie-App im Jahr 2021 gerät? Was kann sie dort beitragen? Wie sehen die Zukunftsperspektiven aus?

So ergibt sich der Aufbau dieses Artikels. Zuerst wird der heutige Stand des Lauten Denkens geschildert, dann richtet sich der Blick auf Thinkie. Die App wird zunächst beschrieben. Am Ende folgt ein Fazit: Wie ist der Stand im Feld? Wie kann es sich weiterentwickeln? Was trägt Thinkie bei?

Seit Ericsson & Simon 1980 hat sich das Laute Denken in seinen Grundannahmen bewährt, aber es hat at im Umfang zugelegt und Leistungen aus dem Umfeld eingemeindet:

Das Laute Denken nutzt den aktuellen Stand der Technik:
- Audio- und Videoaufzeichnungen sind normal. Blickverfolgung kann hinzukommen.
- Internet und Word Wide Web mit ihren integrierten Anwendungen sind von überall her zugänglich.
- Verfahren zur automatischen Spracherkennung sind deutlich anwendungsreifer geworden.

Es erweitert sein Methodenrepertoire und passt sich an unterschiedliche Fachgebiete an:
- Lautes Denken wird oft mit anderen Verfahren kombiniert, um einen Denkprozess zu verfolgen.
- Vielfach wurden umfangreichere Datenmodelle ausgearbeitet. Sie bilden die empirischen Ergebnisse des Lauten Denkens auf die Fragestellung der Studie ab.
- Die Transkription der Verbalprotokolle und die Datenanalyse gehen weit über das hinaus, was Ericsson & Simon vorschlugen.
- Wirtschaftlichkeit ist beim Lauten Denken ein Punkt. Vor allem im Usability Engineering wird die Auswertung der Verbalprotokolle auf das notwendige Minimum reduziert. Die Transkription entfällt ganz oder teilweise, ebenso die Interpretation. Man wertet aus, was Beobachter notiert haben.
- Videos dokumentieren, was die Testperson tut. Man kann aus Worten und Handeln auf die Denkleistung schließen. Die Analyse wird genauer und sicherer.
- Vom Ton der Anfangszeit geht das Laute Denken in eine Multimedia-Umgebung über.
- Benutzererfahrungen spielen heute eine ungleich größere Rolle als 1980. Wenn es darauf ankommt, nimmt man Lautes Denken in der natürlichen Umgebung auf.

Im Anschluss an die Schilderung des aktuellen Umfeldes wird Thinkie 3 dargestellt:

- Thinkie ist eine mobile App für die Datenaufnahme und Transkription beim Lauten Denken.
- Thinkie läuft weltweit im IOS. braucht ein iPhone und iPad mit iCloud-Zugang,
- Der Ton für das Verbalprotokoll wird auf dem iPhone aufgenommen, der Video auf dem iPad.
- Alle Daten werden in einem benutzereigenen Container in der Cloud gespeichert.
- Alle Dateien können in die eigene Systemumgebung exportiert werden.
- Auf dem iPad kann man die Tondatei für eine manuelle Transkription segmentieren oder stückweise zur Transkription zum Server der Speech-To-Text API von Apple schicken. Das Transkript kann man korrigieren.
- Vergleichbare Apps für das Laute Denken stehen im IOS noch aus. Thinkie könnte unter anderem mehr Tools für Transkription und Dateninterpretation anbieten. Thinkie 4 könnte Sinn machen.

Lautes Denken - eine Sicht von heute (2021)

Grundinformationen über den Umgang mit Lautem Denken findet man nach wie vor bei Someren et al. 1994.

Das Arbeitsgedächtnis

Das Denken spielt sich im Gehirn im Gedächtnissystem ab. Das Gedächtnismodell von Atkinson & Shiffrin 1968 ist in großen Zügen bis heute in Gebrauch.

Es postuliert einen Aufnahmebereich, der Sinnesdaten kurzfristig speichert und wiederholt.

Das Arbeitsgedächtnis (Short-term memory - STM) mit geringer Kapazität ist der zentrale Ort des Denkens (Baddeley 2003). Dort werden Inhalte (Chunks) sehr schnell von hinzukommenden Chunks überschrieben und somit vergessen. Im Langzeitgedächtnis (Long-term memory - LTM) wird viel Information auf lange Zeit gespeichert. Wissen aus dem Arbeitsgedächtnis wird dorthin geschrieben.

Ebenso wird beim Denken laufend Material aus dem Langzeitgedächtnis ins Arbeitsgedächtnis geholt. Daneben gehen sensorische Daten in die Denkprozesse ein.

Den inneren Aufbau des Arbeitsgedächtnisses und seine Verbindung mit dem Langzeitgedächtnis illustriert die Abb.1. Es gibt dazu auch andere Theorien (Cowan 2010, Quak et al. 2015). Eine genauere Darstellung bieten Amin & Malik 2013.

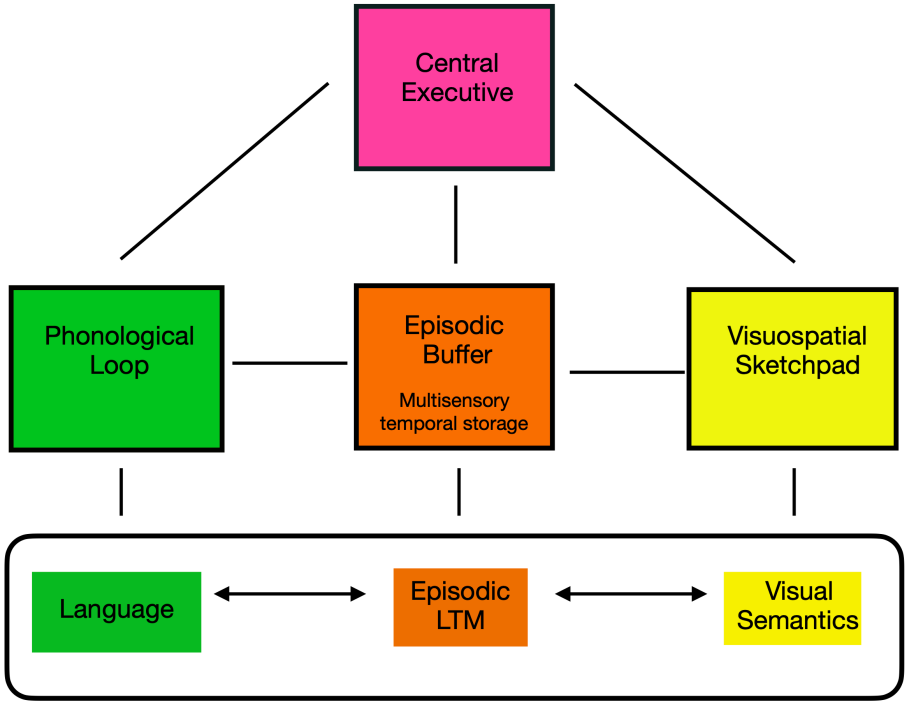

Abb. 1. Arbeitsgedächtnis mit Komponenten von Baddeley, nach Park & Gooding 2014

Denken und Verbalisierung: Lautes Denken

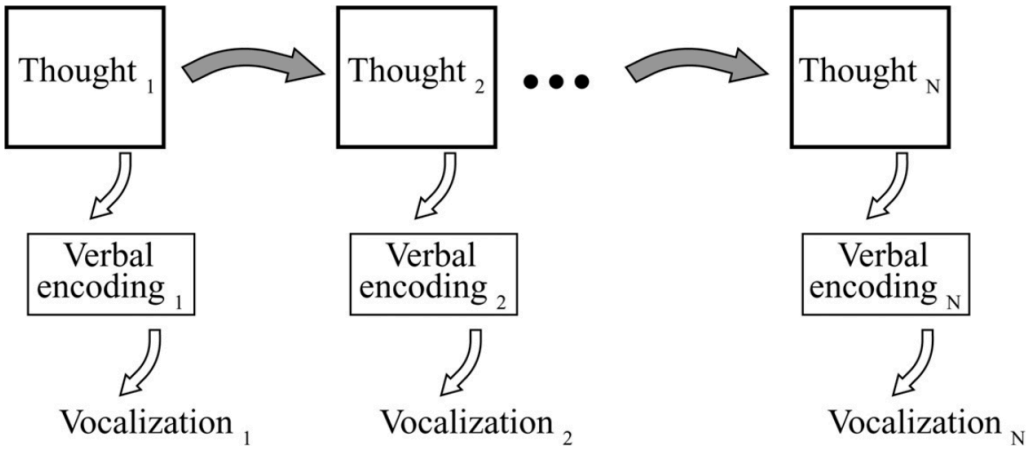

Abb. 2. Denkprozess mit Verbalisierung der Gedanken / Denkschritte aus Ericsson 2006

Wenn ein methodischer Ablauf zu verfolgen ist, folgt eine Denkleistung - ein Gedanke, ein Denkschritt - der anderen. Wenn dabei laut gedacht wird, werden die einzelnen Schritte beim Denken sofort verbalisiert, also von Äußerungen begleitet. Dieses Modell des Denkprozesses und des Lauten Denkens stammt von Ericsson & Simon 1980. Die Abb. 2 von Ericsson 2006 stellt es grafisch dar.

Das Mitsprechen beim Überlegen kommt im Alltagsleben vor, besonders bei Leistungen, die in mehreren Schritten zu erledigen sind, etwa wenn mehrere Schalter eines Gerätes bedient werden müssen, bevor man es starten kann.

Wenn eine Testperson bei einer strukturierten Denkleistung kontinuierlich mitspricht („verbalisiert“), ist das Lautes Denken. Es entsteht ein Verbalprotokoll. Man erhält eine direkte Abbildung des Denkprozesses. Wiedergegeben werden Inhalte des Arbeitsgedächtnisses. Was gesagt wird, ist aktuell richtig, aber nicht vollständig. Es handelt sich um Daten: „Verbal Data is Data“.

Der Auszug aus dem aktuellen Inhalt des Arbeitsgedächtnisses kann Lücken haben:
- Denkleistungen fehlen, wenn sie so automatisiert sind, dass sie nicht bewusst werden.
- Inhalte werden vergessen, insbesondere wenn sie schnell von neuen Chunks überlagert werden.
- Sprechen ist langsamer als Denken. Denkinhalte können auf der Strecke bleiben, weil die Zeit für ihre Verbalisierung fehlt.

Weil bei einem kookkurrenten / synchronen (gleichzeitigen) Verbalprotokoll im Rhythmus des Denkprozesses mitgesprochen wird, verlangsamt sich tendenziell der gesamte Vorgang. Seine Struktur bleibt jedoch erhalten.

Wenn während einer Studie ein Verbalprotokoll aufgenommen wird, ist sehr oft ein Betreuer anwesend (Abb. 3). Sie oder er erinnert die Testperson daran, dass sie reden möge („keep talking“), wenn der Redefluss aussetzt. Ansonsten soll die Testperson nach vorherrschender Auffassung so wenig wie möglich beeinflusst werden.

Abb. 3. Beteiligte beim Usability Test / Thinking Aloud aus Moran 2019

Wenn eine mechanische Aufzeichnung des Verbalprotokolls nicht möglich oder nicht sinnvoll ist, kann ein Beobachter wichtige Äußerungen notieren.

Soweit die theoretische Sicht.

Boren & Ramey 2000 haben die praktische Umsetzung im Usability Engineering untersucht und festgestellt, dass dort die Vorstellungen weit auseinander gehen. Bernardini 2001 beobachtete bei der Anwendung des Lauten Denkens in der Übersetzungswissenschaft ähnliche methodische Differenzen oder Defizite.

Hughes & Parkes 2003 berichten über den Stand im Software Engineering.

Denken ist situationsgebunden

Denken ist situationsabhängig, schon weil Dinge im Umfeld einer Person in ihr Begriffe und größere Zusammenhänge wachrufen:

- Das Stoppschild an der Kreuzung löst ein Verkehrsverhalten aus: anhalten, nachsehen, wenn niemand kommt, weiterfahren, sonst warten, bis die Kreuzung frei ist.
- Ein Tango erinnert an Argentinien und man tanzt ein paar Schritte.

Für Situationen des Lauten Denkens gilt das auch. Eine Testperson, die eine Benutzeroberfläche steuern soll, ist darauf angewiesen, dass sie das Untersuchungsobjekt vor sich hat und es mit Tasten, Stimme oder sonst wie bedienen kann. Ebenso ist es, wenn sie etwas anderes wahrnehmen oder handhaben muss, um ihre Aufgabe zu erledigen, etwa eine Telefonliste, ein Wörterbuch oder einen Notizzettel. Weil sie da sind und benutzt werden, beeinflussen sie das Denken.

Störungen (Lärm…) aus dem Umfeld sollten beim Lauten Denken nicht vorkommen. Ganz allgemein gesehen denkt es sich besser, wenn man sich in einer Umgebung wohlfühlt.

Nach heutigem Stand nimmt man fast immer auch einen Video auf, um die sichtbare und hörbare Umgebung während des Lauten Denkens zu dokumentieren.

Wenn man beobachten kann, wie die Person mit Dingen hantiert, erhält man zusätzliche Hinweise auf ihr Denkverhalten. Was die Testperson in ihrer Umgebung beachtet oder benutzt (die „affordances“ - Gibson 1986), wird verstärkt in die Auswertung einbezogen.

Man kann aus dem Verbalprotokoll und der Video-Darstellung rekonstruieren, was gedacht und gemacht wurde (Triangulation). Dadurch erhöht sich die Genauigkeit und die Sicherheit der Interpretation.

Laborstudien und Feldstudien

Sehr häufig werden Studien des Lauten Denkens im Labor durchgeführt, also dort, wo die beteiligten Forscher arbeiten und die nötige Infrastruktur haben und kennen. Auch finanzielle Gründe können gerade in Industriestudien andere Lösungen ausschließen.

Im Usability Engineering sind Laborstudien in der Überzahl (Nielsen 1994, Fan et al. 2020b, McDonald et al. 2012). Eine begrenzte Validität nimmt man in Kauf (Kieldkov et al. 2004, Joe et al. 2015). Man findet viele, aber nicht alle Benutzungsprobleme. Feldstudien ausserhalb des Labors macht man, wenn sie wirklich nötig sind (Kjeldkov & Skov 2014).

Vorgänge, die nur in einer bestimmten Umgebung möglich sind, zeichnet man dort auf: etwa am Krankenbett (Aitken & Mardegan 2000, Han et al. 2007), in der Krebs-Nachsorge (Jaspers et al. 2009), im fahrenden Auto (Monsalve et al. 2020), während eines Virtual Reality - Erdbebens (Feng et al. 2020) oder im Studio eines Desig A video displays the physical behavior of the test subject. ners (Pringle & Sowden 2017).

Wenn man Protokolle des Lauten Denkens detailliert auswerten will, sollten die Umgebungsparameter stimmen. In der natürlichen Umgebung erhält man eine realitätsnähere Sicht auf die untersuchten Denkvorgänge. Wenn es darauf ankommt, arbeitet man mit den Testsubjekten im Feld, also in deren natürlicher Umgebung. Feldstudien sind aus Forschersicht

Kookkurrent / synchron oder retrospektiv?

Alternativ oder zusätzlich zum kookkurenten / synchronen Verbalprotokoll während der Denkleistung kann man auch im Anschluss daran nach einem retrospektiven Protokoll fragen. Dazu werden die Erinnerungen an den Vorgang aus dem Langzeitgedächtnis geholt. Das kann zu größeren Lücken führen oder zu Uminterpretationen des Denkvorgangs.

Oft erbittet man auch von der Testperson eine gesonderte Interpretation des Geschehens, mündlich oder mit einem Fragebogen. Ebenfalls gängig ist es, das Verbalprotokoll mit der Testperson zu diskutieren. Dabei lassen sich Unklarheiten beheben.

Einen Vergleich der Verfahren findet man bei Alhadreti & Mayhew 2018. Die Autoren plädieren für das weniger zeitaufwendige kookkurrente Verbalprotokoll. Ein früherer Vergleich unterschiedlicher Aufnahme-Modi findet sich bei Olmsted-Hawala et al. 2010.

Datenmodelle

Auf die Datenaufnahme folgt deren Interpretation und Analyse. Bei Ericsson & Simon 1980 findet man dazu nicht viel: auf Abb. 2 folgt ein Gedanke dem anderen. Es bleibt verborgen, was die Gedanken beinhalten und bedeuten.

Wenn jemand sagt “Ich weiß nicht weiter.“ und ein anderer „Rechts herum, auf geht’s!“ unterscheiden sich beide Äußerungen in ihrem kognitiven Gehalt und in den Handlungen, die sie begleiten oder auslösen. Aufgrund dieser Eigenschaften kommen sie in verschiedene Klassen. Das erste Beispiel könnte bei Selbstkontrolle oder Metakognition untergebracht werden, das zweite bei „Start von Handlungen“. Weiterhin enthält das erste Beispiel nur eine Einheit: Eine Wissenslücke wird attestiert. Das zweite gibt jedoch zwei Denkakte wieder: „rechts herum“ (Erkennen der Sachlage) und „auf geht’s“ (Start einer Handlung).

Einzelne Denkakte lassen sich aufgrund ihrer Eigenschaften klassifizieren. Intern können sie zusammengesetzt sein. Sie können Strukturen von mehr als zwei Untereinheiten umfassen.

In einer Studie des Lauten Denkens werden die Denkakte den Untersuchungszielen entsprechend modelliert. Die Datenmodelle sind also in ihrem Anwendungsgebiet theoretisch begründet.

Im Usabilty Engineering bieten sich die Bedienelemente (Starttaste, Exit etc.) auf einer Benutzeroberfläche als erste Ordnungskriterien für Denkakte an. Wo die Testbenutzer Probleme haben, kann man sie nach den Ursachen ordnen und/oder dem Zeiteinsatz.

Die Modelle können vorgeben sein, aber auch aus den Beobachtungsdaten entwickelt werden. Bei Endres-Niggemeyer 1998 sind kognitive Agenten beim Textzusammenfassen zuerst aus einem Selbsttest der Forscherin und dann aus Arbeitsprozessen von sechs Kolleginnen mit Lautem Denken entstanden. Die Methoden zur Datenmodellierung wurden aus dem Knowledge Engineering übernommen (Schreiber et al. 1993, Fensel et al. 1998).

Endres-Niggemeyer 2000 beschreibt eine Demo-Implementierung, in der die kognitiven Agenten unterschiedlicher Funktionsklassen beim Textzusammenfassen von Insektentypen dargestellt werden. Beispielsweise werden die Informationssammler als Bienen dargestellt.

Bei der Entscheidung, ob Facebook - Meldungen falsch oder richtig sind, findet Freilich 2019 bei ihren Probanden drei Hauptstrategien:
- searching for more
- knowledge carries the most weight
- every detail needs to fit

Die Strategien wurden teils deduktiv definiert und von Vorbildern übernommen, aber auch induktiv aus den Beobachtungsdaten entwickelt. Die Verbalprotokolle zeigten auch, dass sich in sozialen Netzwerken oft Meinungen vorfinden, die sich nicht direkt als richtig oder falsch einordnen lassen. Zudem wurden Sekundärkriterien herangezogen wie ein korrektes Zitierverhalten.

Zhou & Lin 2012 untersuchen mit einem Strategienmodell, wie Studierende Texte vom Englischen ins Chinesische und umgekehrt übersetzen. Das Modell entnehmen sie der Literatur und erweitern es für ihre eigene Studie.

Transkription ?

Ob und wie weit man Verbalprotokolle transkribieren muss, ist umstritten. Transkribieren ist arbeitsaufwendig. Was nötig ist, hängt weitgehend vom Gegenstand der Untersuchung ab. Weil mit der Transkription schon verifiziert und wiedergegeben wird, was die Tondatei enthält, beginnt damit auch die Interpretation der Daten.

Nach Nielsen 1994 (und später) bleibt die arbeitsaufwendige Transkription im Usability Engineering weg. Beim „simplified thinking aloud“ notiert ein Betreuer die Probleme, die die Testperson beim Umgang mit einer Benutzeroberfläche verbalisiert. Diese Notizen reichen für die weitere Bearbeitung aus.

Kjeldskov et al. 2004 führen diese Auffassung als „Instant Data Analysis“ weiter. Dabei folgen der Testperson während des Lauten Denkens sowohl ein Betreuer als auch Beobachter in einem anderen Raum. Am Ende tragen alle Beteiligten ihre Beobachtungen in einem Verfahren der Ideengenerierung („Instant Data Analysis - IDA“) zusammen. Weil mehrere Beteiligte kooperieren, werden die meisten Schwachstellen von Interfaces erfasst. Die Probleme werden nach Schwere klassifiziert.

Joe et al. 2015 verwenden IDA bei der Untersuchung einer medizinischen Benutzerplattform für ältere Patienten. Auch sie erreichen befriedigende Ergebnisse, indem sie Testbenutzer die Benutzerplattform mit Lautem Denken anwenden lassen und im Anschluss eine gemeinschaftliche Ideengenerierung mit Pinnwand und Kärtchen durchführen. Dabei werden die Probleme von der Gruppe benannt und im Gruppenkonsens klassifiziert.

Wenn man nur untersuchen will, wie in bestimmten Konstellationen reagiert wird (Gefühlsäußerungen, Störungen, Bedienfehler), reicht es eventuell aus, nur Teilpassagen aus dem gesprochenen Verbalprotokoll zu transkribieren und zu analysieren. Man gibt dann deren Start- und Endzeiten im Transkript an.

Feng et al. 2020 untersuchen, wie sich Personen in einem Immersive Virtual Reality - Erdbeben in einem Krankenhaus in Auckland verhalten. Die Verbalprotokolle werden von der VR-Brille aufgenommen. Der Audio wird mit dem Video der VR-Episode synchronisiert und in einer MP4-Datei gespeichert. Ein einfaches Kodierschema hält zunächst Situationen fest (z.B. während des Erdbebens oder danach). Dazu kommen typisierte Aussagen über das, was die Testpersonen tun und wie sie sich über die Sachlage informieren. Sie laufen im Treppenhaus etwa hinter anderen Personen her oder prüfen nach, ob der Aufzug funktioniert. Rater kodieren Vorkommen solcher Situationen und Verhalten.

Sakar & Chakrabarti 2007 haben Kollegen mit Transkriptionserfahrung nach häufig auftretenden Problemen befragt. Dafür haben sie technische Lösungen gesucht. Ihr Tool zur Unterstützung der manuellen Transkription halbiert die Zeit, die für eine Transkription gebraucht wird.

Analysetools wie ODCS (MacLin & MacLin 2005) und CAPAS (Crutcher 2007) oder kommerzielle Angebote wie MAXQDA und andere arbeiten mit unterschiedlichen Eingabedaten, also Tondaten, schriftlichen Verbalprotokollen, numerischen Werten usw.

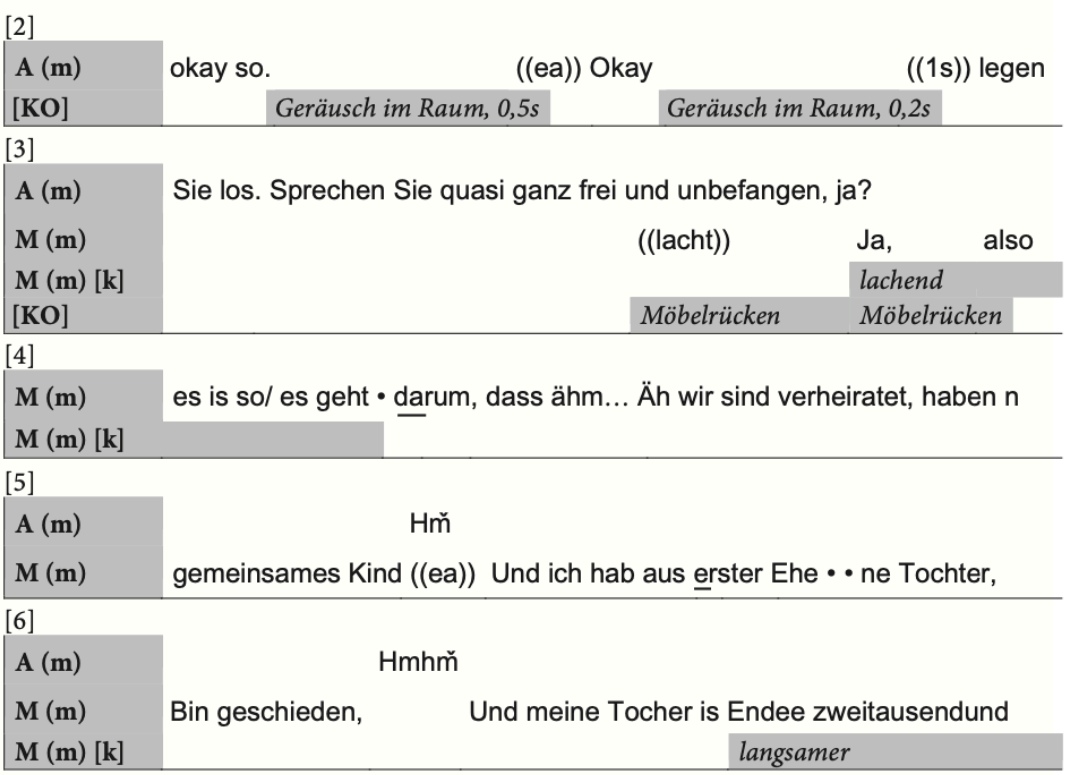

Abb. 4. Transkript in Partiturform: Auszug mit Spur für Anwalt, Mandant und Hintergrund aus Pick 2015

Wenn man das mündliche Verbalprotokoll transkribiert, geht dies oft über eine einfache Umsetzung von gehörtem in geschriebenen Text hinaus. So kodieren Hutchison et al. 2015 auch Pausen und Gesten (im konkreten Fall der Informationssuche), um so den aktuellen kognitiven Prozess besser nachzuvollziehen.

Was die Testperson gesagt hat, ist nicht immer leicht auszumachen. Wer es niederschreibt, muss manchmal mehrfach hinhören, selbst interpretieren und kommt nicht immer zu einem Schluß, sodass eine unklare Stelle bleibt und notiert werden muss. Die Regeln für Orthografie und Zeichensetzung hört man nicht, aber man hält sie ein, soweit es möglich und sinnvoll ist.

Es kommt nicht nur auf den Wortlaut an. Selbst bei einer einfachen Transkription notiert man Pausen, Interjektionen („ahh“, „hmm“, „soso“ …), Änderungen der Lautstärke, die Äußerung von Gemütslagen wie Lachen oder Seufzen. Sie sind wichtig für die Abgrenzung von Denk- und Aktionsschritten

Für anspruchsvollere Transkripte eignet sich eine Partiturschreibweise, die parallel festhält, was auf den einzelnen Kanälen zu beobachten ist (Abb. 4). Anwalt und Klient haben hier zusätzlich zur gesprochenen Äußerung eine Spur für den Sprechmodus („lachend“, „langsamer“). Die einzelnen Events können Zusatzangaben wie Start- und Endzeiten oder Klassifikationskodes mitführen. Umgebungsgeräusche haben ebenfalls eine Spur: „Möbelrücken“.

Für die manuelle Transkription gibt es Anleitungen im Netz, die auf unterschiedliche Anforderungen und Systemumgebungen eingehen: Klein 2015, Setting et al. 1998, Rincon 2018, Dresing et al. 2015, oder auf Plattformen und Tools wie media-ai oder audiotranscription.de . Dort wird auch eine automatische Transkription angeboten.

Wo eine automatische Transkription möglich ist, kann man sie integrieren. Den Text, den sie liefert, muss man oft korrigieren. Zusätzliche Angaben von der Tonhöhe bis zu Hintergrundgeräuschen muss man bei Bedarf nachtragen.

Eixterne Anbieter (etwa NVIVO , Amberscript oder Otter) können die gesamte Transkription übernehmen.

Datenanalyse

Bei der Datenanalyse wird ein Bestand von Verbalprotokollen auf Datenwerte untersucht. Was dabei passiert, richtet nach dem Untersuchungsgegenstand und dem Grundaufbau der Studie.

Eine Studie kann mehr Eingabekanäle haben als ein Verbalprotokoll, das eventuell Begleitdaten wie eine Blickverfolgung oder einen Video mitführt. Hinzu kommen kann beispielsweise ein Fragebogen oder ein nachträgliches Interview mit der Testperson oder auch eine Kostenanalyse. Sie gehen dann in einer geeigneten Form in das Gesamtergebnis ein.

Das Untersuchungsziel kann sehr eingeschränkt sein, etwa wenn es nur einfache Werte wie den Zeitbedarf für Operationen auf einer Webseite verzeichnet. Es kann aber auch viel umfangreichere Daten verlangen, etwa wenn es fragt, welche Webseiten der Testbenutzer mit welchem Erfolg zur Beantwortung einer Frage herangezogen hat.

Eine quantitative Analyse mit statistischen Verfahren zielt auf numerische Werte wie den Zeitbedarf oder die Anzahl der Versuche. Da Verbalprotokolle gesprochener oder geschriebener Text sind, werden sie jedoch meistens eine qualitative Analyse erfordern. Sie kann mit einer quantitativen Analyse kombiniert werden.

Methodisch kann man einen deduktiven Ansatz mit vorgegeben Fragen verfolgen oder einen induktiven, der die Kategorien der Untersuchung aus den Beobachtungsdaten abstrahiert. Mischformen gibt es auch.

Methodenwissen zur qualitativen Datenanalyse findet man bei NN 2014?, 2019, Lester et al. 2020, mit Bezug auf Unterstützungssysteme bei Gizzi & Rädiker 2021, für MaxQDa und für NVIVO bei Jackson & Bazeley 2019. PAT Research bietet einen Review verfügbarer Systeme zur qualitativen Datenanalyse an. Dort erfährt man von anderen Anbietern.

Eine kombinierte Methodik zur qualitativen und quantitativen Untersuchung der Benutzung von mobilen Karten im Feld und im Labor schlagen Elzakker et al. 2008 vor. Sie machen die besten Erfahrungen mit einer Kombination von Videoaufzeichnungen, Screen-Logging, synchronem Lautem Denken und einer halbstrukturierten Befragung der Benutzer.

Je nach Untersuchungsziel können die Methoden in einem breiten Bereich streuen. Ein paar Beispiele dafür:

- Kreatives Denken beim Gartendesign (Pringle & Sowden 2017)
- Regeln für autonomes Fahren (Monsalve et al. 2020)
- Entscheidungen bei Erdbeben im Krankenhaus (Feng et al. 2020)
- Interpretation von Röntgenaufnahmen (Yoon et al. 2020)
- Übersetzen (kritischer Überblick bei Bernardini 2001)
- Methoden im Übersetzungsprozess (Sun 2011)
- Lesen und Verstehen lernen (Wang 2016)
- Verstehen komplexer Texte (Dahl et al. 2021)
- Fremdsprachentests (Green 2007)
- Interkulturelle Unterschiede beim Usability Test von Chinesen und Dänen (Shi 2010)
- App für mehr körperliche Bewegung (Fischer et al. 2019)
- Medizinische Falschinformation bei der Informationssuche (Ghenai et al. 2019)
- Selbsteinschätzung von Patienten (Al-Janabi et al. 2013)
- Krankenpflege in der Intensivmedizin (Han et al. 2007)

Software-Tools für die Datenanalyse

Wenn man mehrere Inputs integriert behandelt, kann die Datenanalyse schwieriger werden. Die Testperson kann etwas sagen (Tonkanal), sie kann gleichzeitig einen Gegenstand bewegen oder etwas notieren (Videokanal), sie kann die Blickrichtung wechseln (EyeTracking - Oh at al. 2013).

1994 beobachten Levy & Ransdell wie Texte geschrieben werden, indem sie den Ton / das Verbalprotokoll und die Anschläge auf der Tastatur aufnehmen und die Daten auf einen Video projizieren. Der wird per TV vorgespielt. Die Teilprozesse der Textproduktion (planning, reviewing, generating text, revising nach Kellogg 1987) werden manuell kodiert und mit Zeitangaben und Frequenzen kombiniert.

Ein Tool zur Darstellung mehrspurigen Inputs ist ELAN (Abb. 5 - Trippas et al. 2017). Man kann Intervalle aus Video und Audio abgreifen und annotieren. Dazu gibt es zusätzliche Hilfen. Die Annotationen können auch ein Transkript enthalten.

Asselin & Moayeri 2010 arbeiten in einem Marktforschungssystem zur Benutzungsverfolgung (Morae) mit gesprochenem Lautem Denken und Video. Sie wollten bei zwei Jugendlichen etwas über deren Internetkompetenz herausfinden. Die beiden erledigten in ihrer eigenen Umgebung Hausaufgaben. Sowohl im Verbalprotokoll wie im Video wurde auf kognitive und auf affektive Reaktionen hin ausgewertet.

Maher et al. 2018 beschreiben, wie sie mit NVIVO und Haftnotizen Designprozesse darstellen. Die Äußerungen, die Lautes Denken ausmachen, können aus ihrer Sicht nicht nur sprachlich kodiert sein, sondern auch haptisch. Abb. 6 zeigt ihren kombinierten Output von NVIVO-Daten und bunten Haftzettelchen.

Ghenai et al. 2019 beobachten, wie die Darstellung im Webbrowser Benutzer beeinflusst, Meldungen für falsch oder richtig zu halten. Sie nehmen einen Video mit Tonkanal auf und

verfolgen den Blick der Testsubjekte (Eye Tracking). Zum synchronen kommt ein retrospektives Protokoll hinzu. Darin hilft die Blickverfolgung, sich richtig an Vorgänge zu erinnern. Kodiert wurden die Beobachtungsdaten mit Software von **NVIVO**.

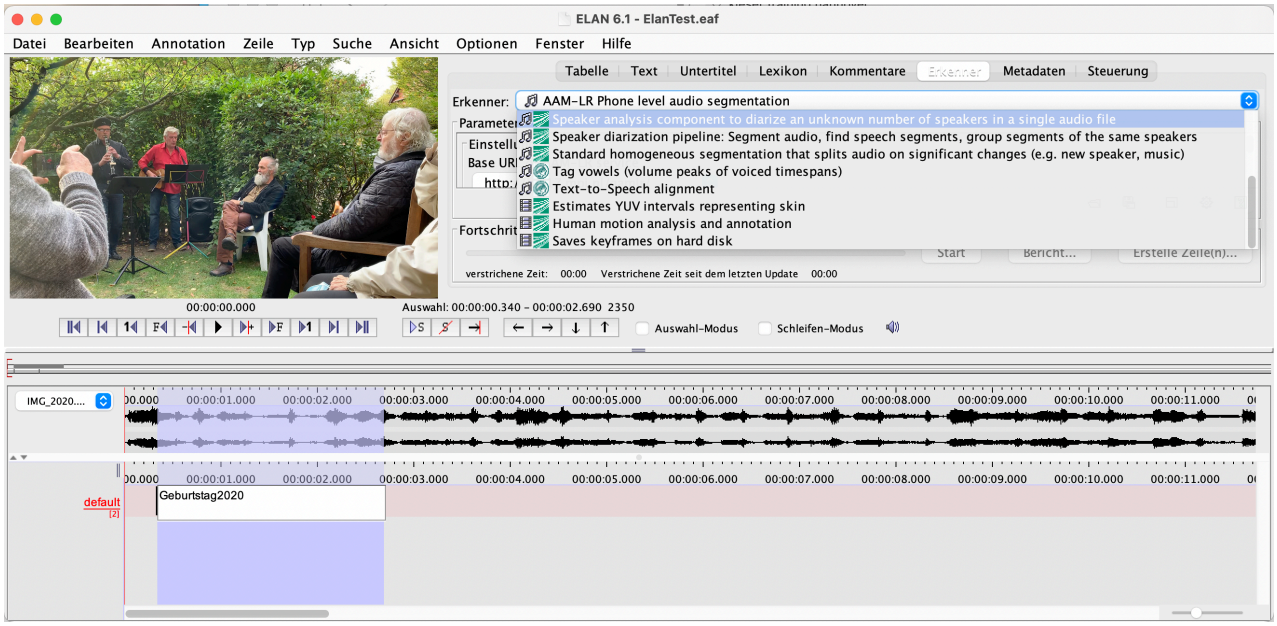

Abb. 5. ELAN: Video und Audio, Annotation (“Geburtstag 2020”) und Erkennungshilfen

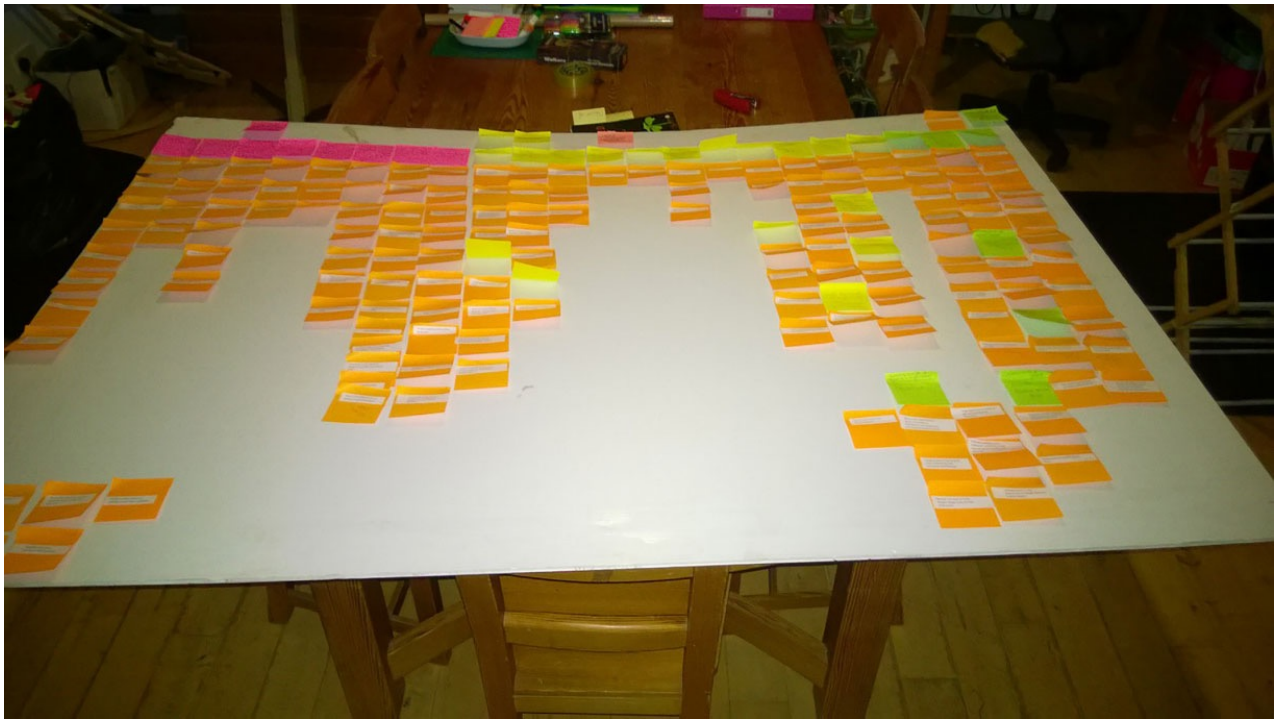

Abb. 6. NVIVO-Output mit Haftzettelchen: kombinierte Darstellung aus Maher et al. 2018

Lautes Denken mit maschinellem Lernen (ML)

Johannsen et al. 2019 finden wie Kjeldskov et al. 2004 und Joe et al. 2015 das traditionelle Laute Denken zu zeitaufwändig und arbeitsintensiv. Ihre Anwendungsumgebung ist Continuous Software Engineering (CSE): dort verteilen Entwickler Software-Verbesserungen kontinuierlich an die Benutzer, sobald sie verfügbar sind..

Die Autoren automatisieren die Datenauswertung mit Verfahren des maschinellen Lernens. Sie verwenden eigene manuelle Transkripte als Trainingsdaten. Den Classifier-Prototypen entwickeln sie ebenfalls selbst. Er klassifiziert nach Gefühlen: positiv, negativ, neutral, unsicher. Darauf wird das Laute Denken der Benutzer untersucht.

Der Classifier liefert in kurzfristig Ergebnisse. Wo Benutzer unsicher oder negativ reagieren, sollte der Entwickler es besser machen.

Goldberg et al. 2020 experimentieren ebenfalls mit maschinellem Lernen. Sie werten Audiomitschnitte von Patienten und Therapeuten aus, um etwas über das therapeutische Bündnis von Patient und Therapeut („alliance“) gegen die Krankheit in einer Therapie zu erfahren. In Lern- und Testdaten bestimmen sie Unigramme / Bigramme mit Sent2Vec und tf-idf .

Fan et al. 2019, 2020a wollen die aufwändige Auswertung von Verbalprotokollen ebenfalls mit ML-Verfahren automatisieren. Sie verwenden 12 transkribierte Thinking-Aloud-Aufnahmen mit zusätzlich ermittelten Merkmalen wie Sentiment oder Pitch. Die Aufnahmen werden Ratern unterbreitet, die Sequenzen von Interesse mit den Kategorien Reading, Procedure, Observation und Explanation kennzeichnen. Eine Problemspanne ergibt sich als Abfolge dieser Kategorien. Nun wird die Zuordnung von Kategorien und Merkmalen zu den Trankriptsequenzen nach Precision, Recall und F-Wert berechnet. Insgesamt ergibt sich, dass die Werte aus der Kombination aller Merkmale besser ausfallen als wenn nur auf der Basis des Transkript gerechnet worden wäre.

Thinkie 3

Eine App im Werden

Thinkie unterstützt auf iPhone und iPad die Datenaufnahme und -bearbeitung in Studien des Lauten Denkens.

Hintergrund ist die Erfahrung der Autorin aus eigenen Feldstudien. Um das professionelle Textzusammenfassen mit Lautem Denken zu erforschen, zog ich in Deutschland und den USA zu einschlägigen Kolleginnen, damals mit einem Laptop, einem mobilen Drucker, einem Audiorecorder und einer Menge Papier - Ergebnis in Endres-Niggemeyer 1998. Günstiger gewesen wäre eine handliche mobile technische Ausrüstung, die überall funktioniert und im Prinzip in die Handtasche passt. Wenn es das noch nicht gibt, war die Überlegung, dann wird es eben gemacht. Auf einem iPhone und einem iPad kann man es doch realisieren.

Es entstand eine erste Thinkie - Version von 2013. Sie stellte das Modell vor (Endres-Niggemeyer 2013).

Thinkie 2 erschien um 2015 herum. Es machte Fortschritte in Richtung Anwendbarkeit und passte sich an das aktuelle IOS an. Es gab mehr Spracherkennung und einen CloudDrive-Zugang.

Thinkie 3 von 2021 passt sich wieder an die aktuelle Systemsoftware an. Die Spracherkennung (Speech-to-Text) ist auf dem aktuellen Stand. Alle Daten werden im eigenen Container des Benutzers in der Cloud gespeichert. Sie lassen sich exportieren. Thinkie 3.1 verlangt eine explizite Autorisierung mit der Apple ID.

Thinkie 3 wird im Folgenden beschrieben. Seine Hauptmerkmale:

- Mobilität: Handtaschenformat auf iPhone und iPad
- Betrieb weltweit im IOS
- iCloud - Verbindung
- Datenspeicherung im benutzereigenen iCloud-Container
- Datenmodell für Cloud-Dateien
- Audioaufnahme auf dem iPhone, parallel Videoaufnahme auf dem iPad
- Bearbeitung / Transkription der Verbalprotokolle auf dem iPad
- Spracherkennung (Speech-To-Text)
- Export der Cloud-Dateien
- Farbenfrohe Benutzeroberfläche

Thinkie 3 auf iPhone und iPad mit Cloud-Anbindung

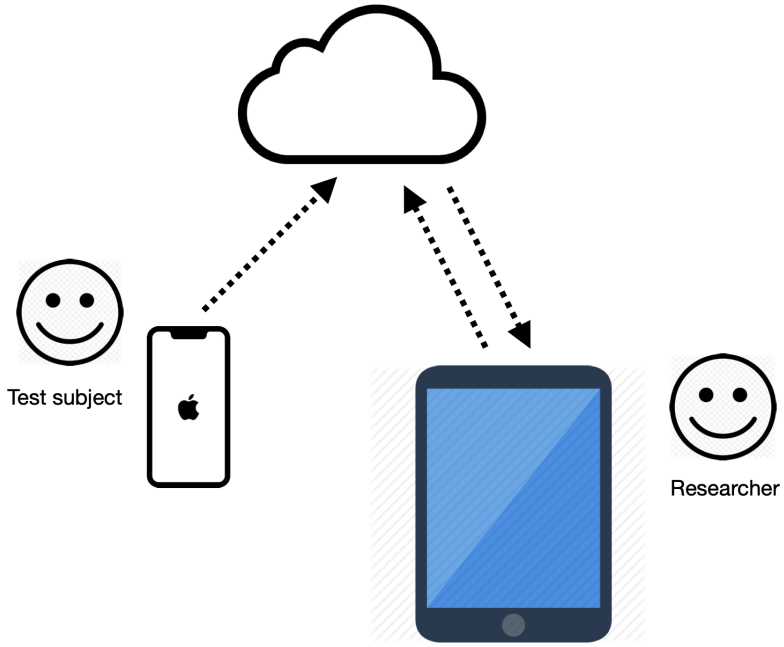

Abb. 7. Mobile Arbeitsumgebung: iPhone für die Testperson, iPad für die Betreuerin / Forscherin, Cloud-Verbindung für die Daten

Um Thinkie zu benutzen, braucht man nicht viel: ein iPhone, ein iPad und eine iCloudverbindung (Abb.7). Das ist die gesamte Systemumgebung. Thinkie funktioniert überall, wo ein ausreichender Internetanschluss vorhanden ist. Das ist vor allem nützlich für die Feldforschung. In Laborstudien schadet es nicht.

Der personelle Aufwand ist ebenfalls gering. Die Testperson benutzt das iPhone und spricht ihr Verbalprotokoll. Die Betreuerin stellt das iPad richtig ein und auf, so dass es die Testperson in ihrem Umfeld abbildet, und sorgt für die Videoaufnahme.

Wenn die Aufnahme beendet ist, schaut / hört man sich die Ergebnisse eventuell an und schickt sie nacheinander zur Cloud-Container. Die weitere Bearbeitung der Aufnahmen erledigt die Betreuerin. auf dem iPad.

Auf dem iPad kann man das Verbalprotokoll und den Video inspizieren. Testperson und Betreuerin können es vor Ort oder sonst wo diskutieren. Die Betreuerin kann das Verbalprotokoll auch in einfacher Form transkribieren und weiterverarbeiten.

Für eine weitergehende Analyse und andere Zwecke kann man die Untersuchungsdaten aus iCloud exportieren.

Autorisierung mit Apple ID

Eine explizite Autorisierung der Thinkie-App mit der AppleID ist die Voraussetzung dafür, dass die Anbindung der App an den privaten iCloud-Container der Benutzerin funktioniert.

Die AppleID wird gleich beim Start der App verlangt. Die Form richtet sich nach den lokalen Gegebenheiten. Das Beispiel in Abb. 8 zeigt eine Fassung mit Gerätecode auf dem iPad. Alternativ könnte auch eine Gesichtserkennung laufen.

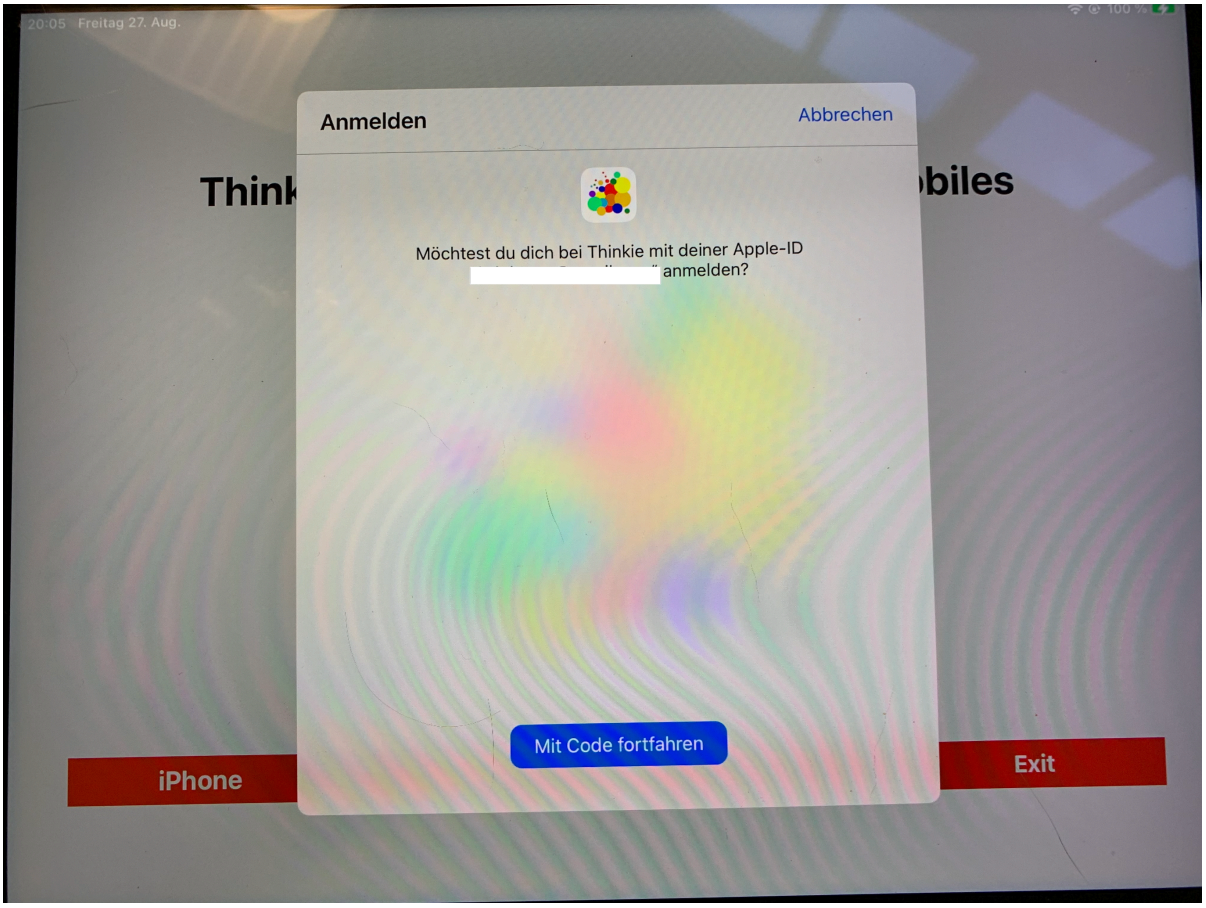

Abb. 8. Autorisierung mit AppleID - hier Fassung mit Gerätecode

ThinkTon: Datenmodell von Thinkie im iCloud-Container

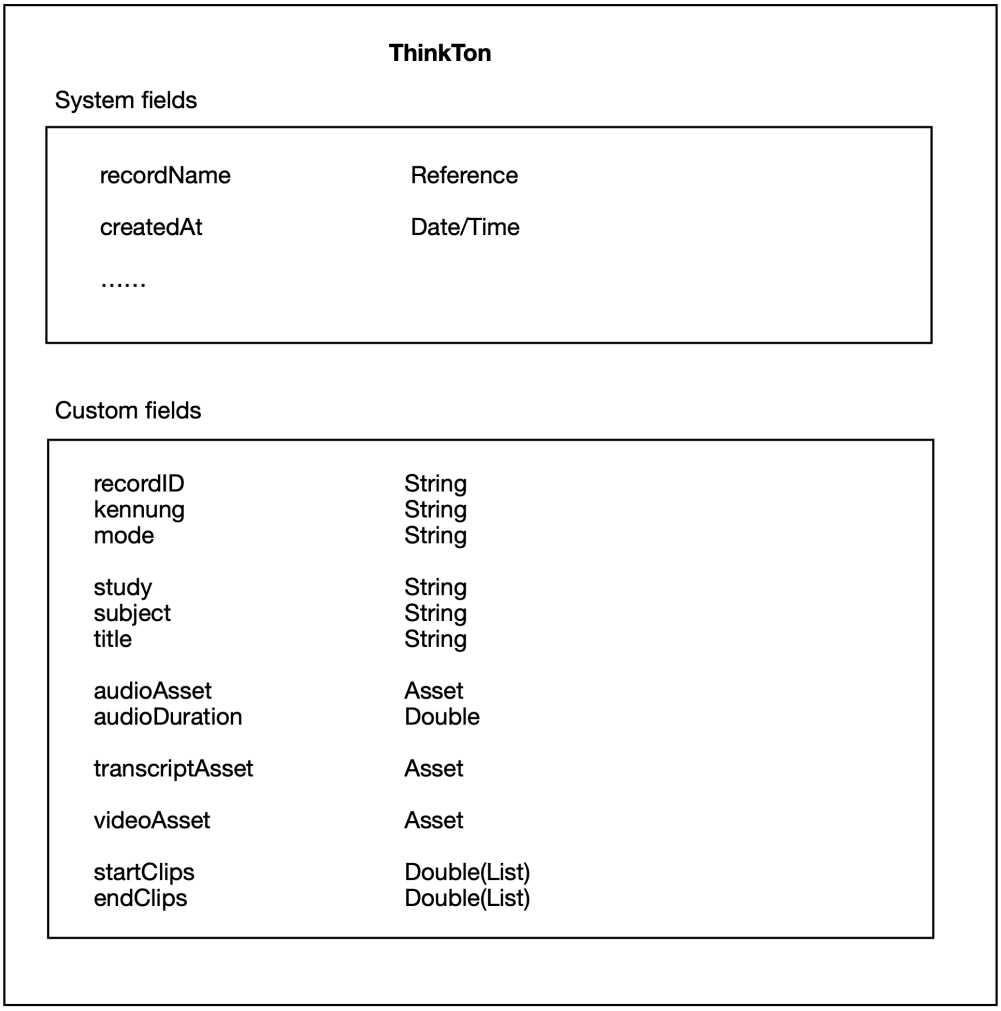

Abb. 9. Das ThinkTon - die zentrale Datenstruktur von Thinkie 3 im benutzereigenen Cloud-Container

Eine Datenstruktur namens ThinkTon (Abb. 9) speichert die permanenten Daten von Thinkie. Sie liegt in der privaten Datenbank des Benutzers. Das iPhone schreibt Sätze des Typs ThinkTon in die Sektion „phone“ der Datenbank, das iPad schreibt in die Sektion „pad“.

Von den systemeigenen Datenfeldern wird nur der Identifikator „recordName“ und das Entstehungsdatum „createdAt“ benutzt.

Im benutzereigenen Bereich steht in „study“, „subject“ und „title“ ein dreiteiliger Namen des Records. Die drei Strings werden zum Datensatznamen „kennung“ verkettet. Damit wird der richtige Record gesucht.

Die Untersuchungsdaten stehen in Assets. Assets nehmen Daten in großen Umfängen auf. Es gibt je einen für Audiodaten, für Videodaten und für Transkriptdaten.

Die Dauer der Audiodatei wird gesondert angegeben, die Dauer der Videodatei wird der Datei selbst entnommen.

Die Listen „startClips“ und „endclips“ enthalten die Anfänge und Enden von Segmenten der Audiodatei, wenn diese für die Transkription in handhabbare Abschnitte untergliedert wurde.

Ein Clip ist ein timestamp des Datentyps Double (“double-precision, floating-point value“).

Thinkies Start von iPhone und iPad

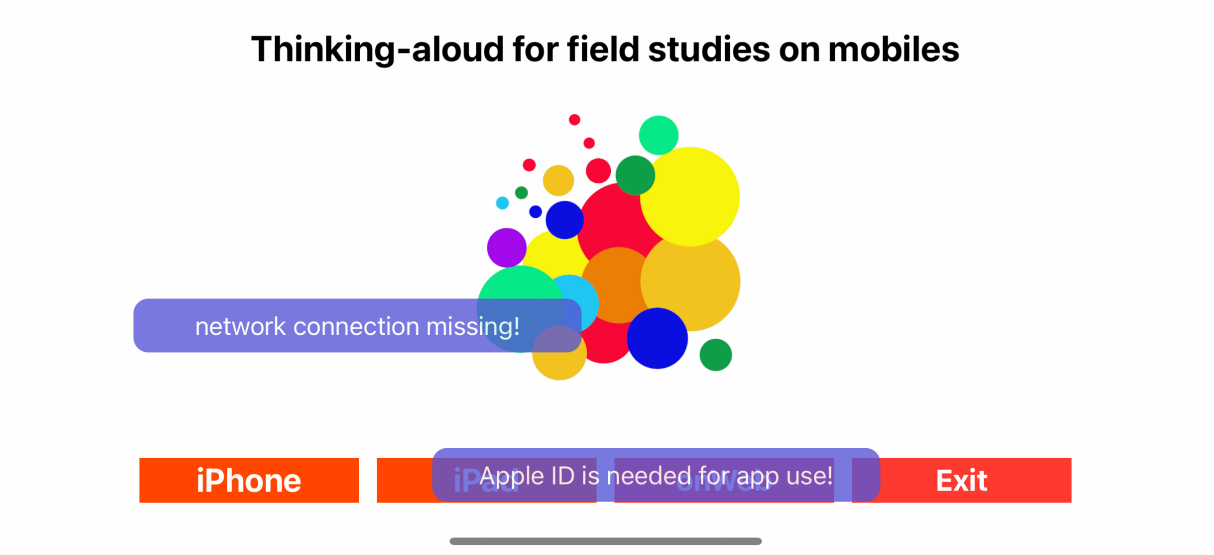

Abb.10. Gemeinsamer Startbildschirm von Thinkie auf iPhone und iPad: WiFi / WLAN und AppleID fehlen

Thinkie meldet sich mit einer gemeinsamen Seite von iPhone und iPad (Abb. 10). Sollte die App nicht betriebsbereit sein wie in Abb.10, erscheinen blaue Warnmeldungen und der Zugang zu iPhone und iPad ist gesperrt. Blaue Meldungen erscheinen auch anderswo, um eine Reaktion der Benutzerin auszulösen.

Mit „onWeb“ findet man Zusatzinformationen, darunter auch deutsche und englische Arbeitsanleitungen für Benutzer.

„Exit“ schaltet zurück zum IOS.

Auf dem iPhone

Auf dem iPhone gibt es nur eine Seite. Ihren Funktionsumfang zeigt Abb.11, die Benutzersicht präsentiert Abb. 12.

Auf dem iPhone wird aufgezeichnet, was die Testperson sagt. Die Sprache ist beliebig. Die Tondatei enthält das gesprochene Verbalprotokoll. Geräusche werden ebenfalls aufgenommen.

Die Tondatei wird mit drei Elementen benannt, die standardmäßig als Studie, Titel und Subjekt kategorisiert sind (Abb. 12). Ihre Belegung ist beliebig. Aus ihnen entsteht der String „kennung“, der den Datensatz identifiziert. Damit findet man die Aufnahme später wieder.

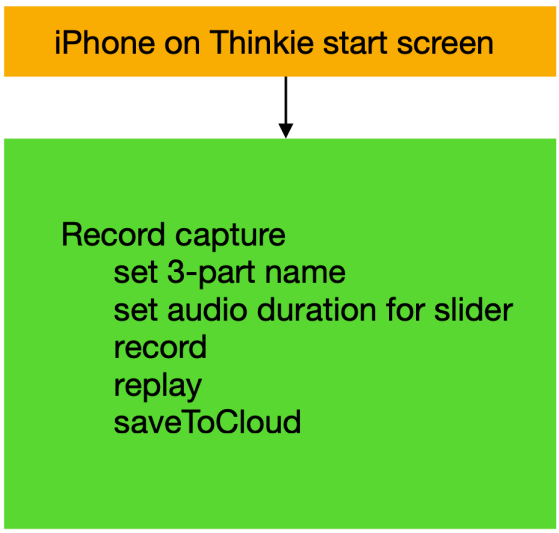

Abb. 11. iPhone-Bildschirm: Funktionen

Ein Schieberegler läuft bei der Aufnahme mit. Die erwartete Dauer der Aufzeichnung kann man eintragen. Dauert die Aufnahme länger, bleibt der Regler stehen, die Aufnahme geht jedoch weiter.

Wenn die Aufnahme beendet ist, kann man sie anhören. Ist sie fehlerhaft, produziert man besser eine neue. Will man sie behalten, sichert man die Audiodatei im iCloud-Container („save“). Ein Erfolg wird mit „stored!“ bestätigt. Alternativ kommt eine Fehlermeldung.

Für die Besitzerin des iPhones ist die Erlaubnis zur Aufnahme voreingestellt. „Recording allowed“ erscheint. Nur beim allerersten Aufruf wird man aufgefordert, dem Gebrauch des Mikrofons zuzustimmen (Abb.13). Ohne Zustimmung geht es nicht weiter.

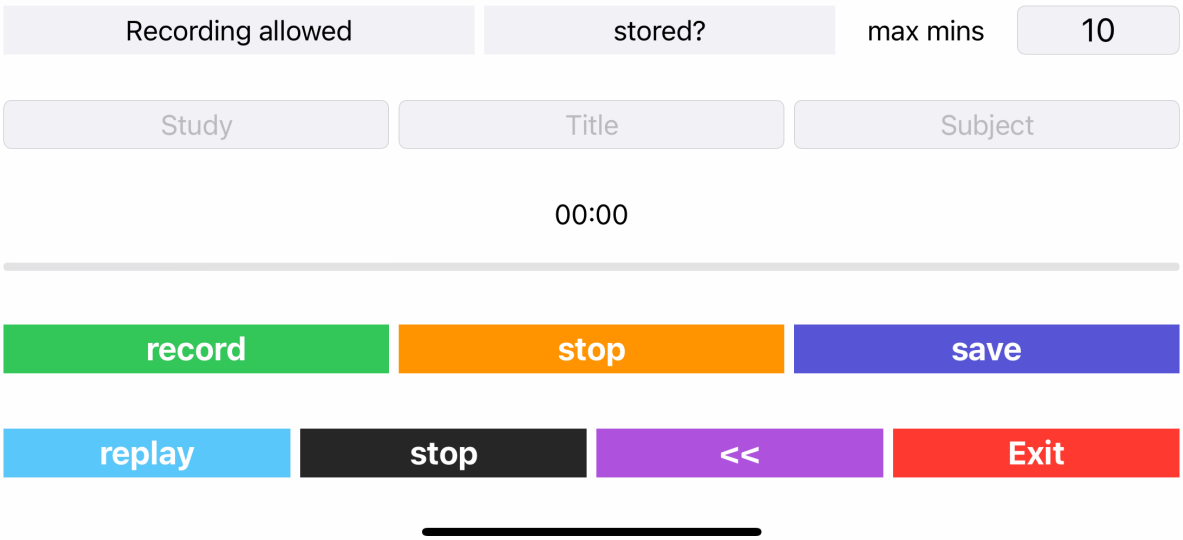

Abb. 12. Bildschirm für die Tonaufnahme im Startzustand

Abb. 13. Bitte um die Erlaubnis des Benutzers, das Mikrofon zu benutzen

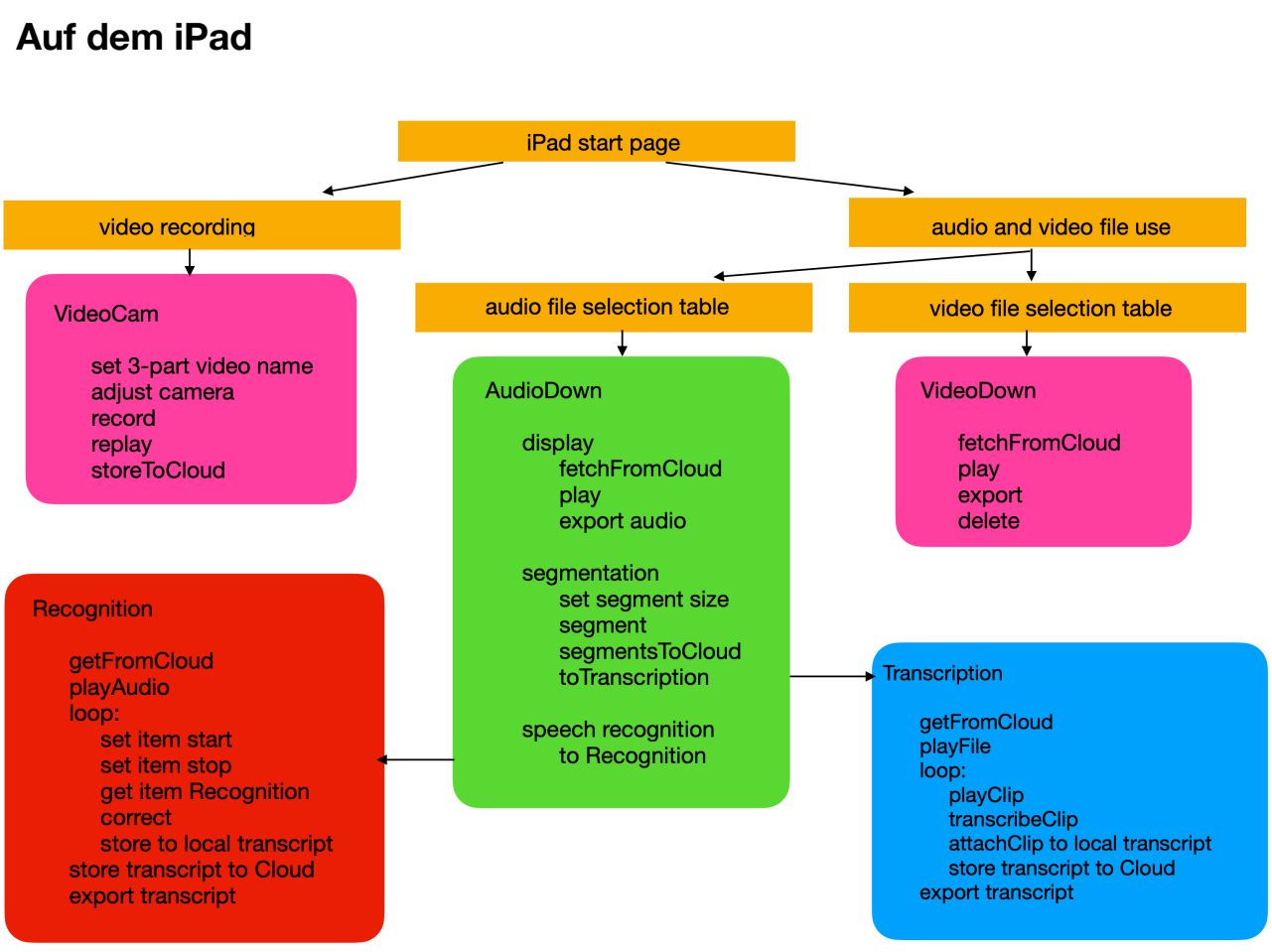

Abb. 14. Funktionsangebot von Thinkie auf dem iPad

Den ersten Auftritt von Thinkie 3 auf dem iPad zeigt Abb. 15.

Von außen gesehen (Abb. 7) erscheinen iPhone und iPad gleichwertig. Betrachtet man jedoch das Funktionsangebot (Abb. 11 und 14), dann wird das iPad ungleich stärker beansprucht als das iPhone.

Für die einzelnen Aufgaben gibt es auf dem iPad Benutzeroberflächen, die man über Verteiler (gelb in Abb.14) anspringen kann. Sie kommunizieren mit dem Cloud-Container. Darin werden Daten gesichert, von dort holt man sie sich wieder.

Videos werden auf der Seite „VideoCam“ (Abb. 17) aufgenommen und im Cloud-Container gespeichert. Zum Abspielen holt man sie aus dem Cloud-Container ab (Seite „VideoDown“ - Abb. 18).

Die Seite zur Tonverarbeitung („AudioDown“ - Abb.19) leitet weiter zu zwei gesonderten Seiten für die manuelle Transkription („Transcription“ - Abb. 20) und die Spracherkennung („Recognition“ - Abb. 21).

Wegen unterschiedlicher Anforderungen folgt die Audiobearbeitung zwei Linien:
- Bei einer Segmentierung für die manuelle Transkription werden die Start- und Endpunkte der Segmente gespeichert. Die Segmente werden auf der Folgeseite „Transcription“ in einer Tabelle zur Bearbeitung angeboten. Den Text schreibt man in das Transkript.
- Auf der Recognition-Seite greift der Benutzer sinnvolle Tonsequenzen ab, die zum Speech-Server geschickt werden. Dessen Output wird an das Transkript angehängt.
- Das aktuelle Transkript holen die Transkription- und die Recognition-Seite aus dem Cloud-ThinkTon ab. Beide können von dort den aktuellen Stand abholen und speichern.

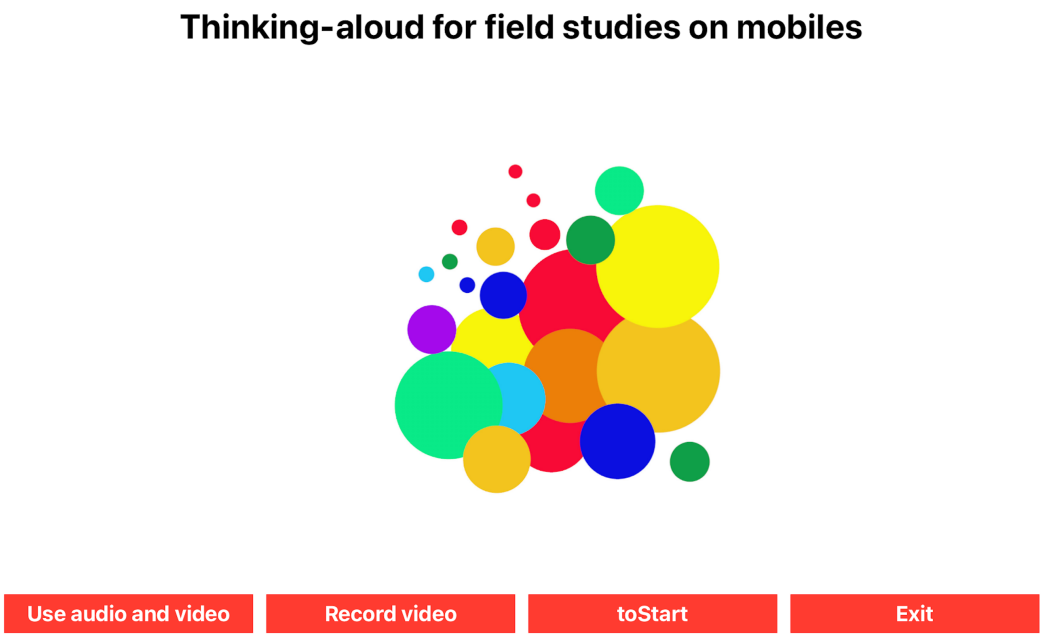

Abb. 15. Startbildschirm auf dem iPad

Vom ersten iPad-Bildschirm (Abb. 15) führt die Taste „Record video“ zur Seite für die Videoaufnahme (Abb.17). Dort gibt man dem Video einen Namen, nimmt ihn auf, spielt ihn eventuell noch einmal ab und sichert ihn im Cloud-Container. Von dort kann man ihn wieder abholen (Abb.18), um ihn abzuspielen oder zu exportieren.

Die Taste „Use audio and video“ führt vom iPad-Startbildschirm aus zu einer Tabelle, die alle verfügbaren Ton- bzw. Videodateien auflistet (Abb. 16). Dort wählt man eine einzelne Audio- oder Video-Datei aus.

Mit der Taste „toStart“ kehrt man zur gemeinsamen ersten Seite von iPhone und iPad zurück, „Exit“ beendet die App schaltet zurück ins IOS.

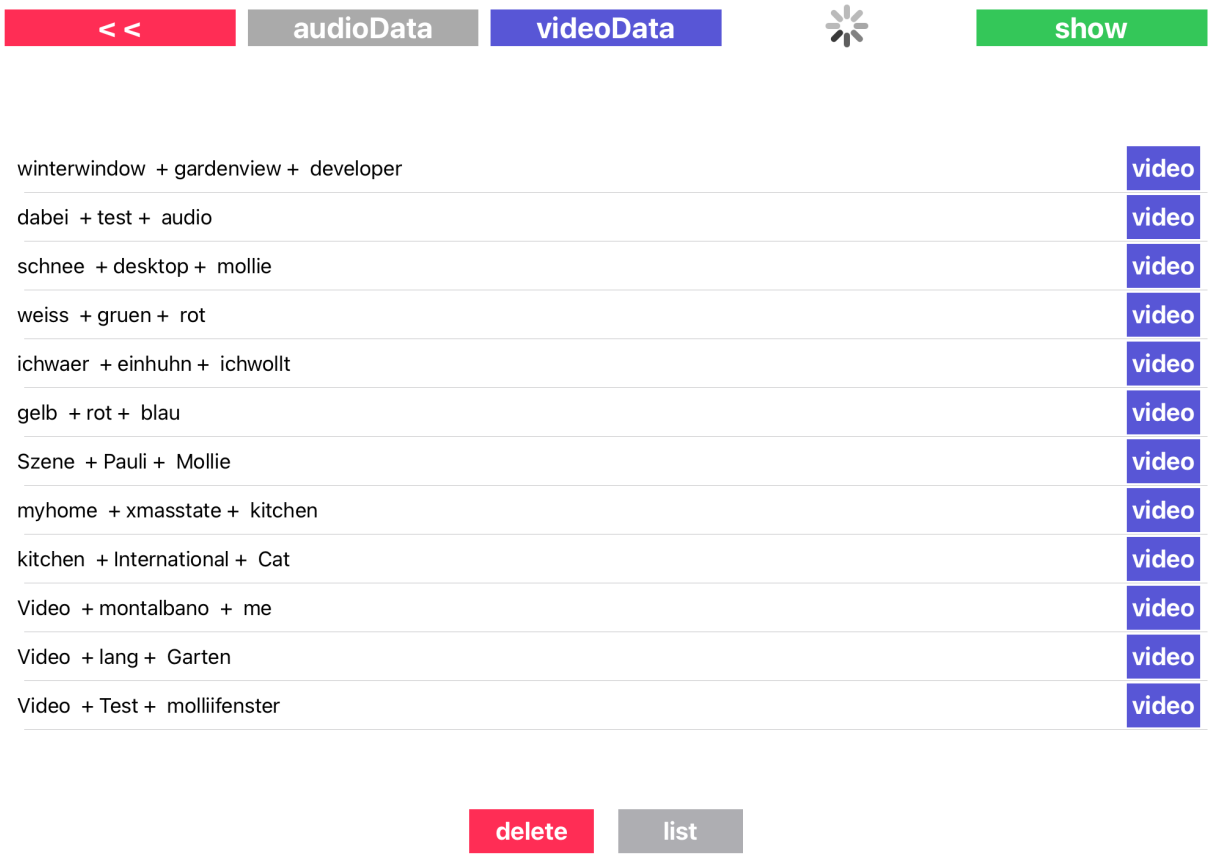

Abb. 16. Tabelle zur Auswahl einzelner Videos / Tondateien mit Einstellung Videos

Videos aufnehmen und abspielen

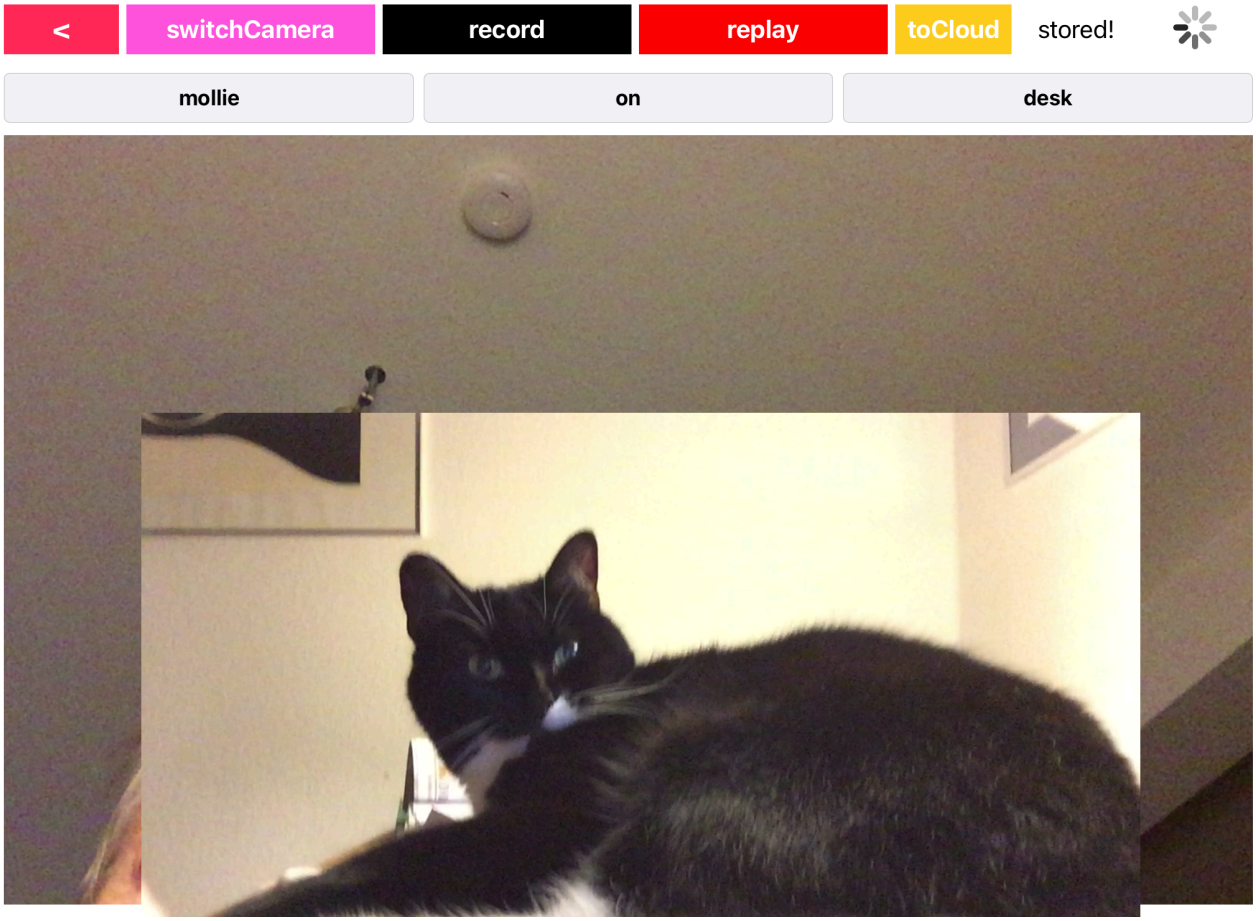

Abb. 17. Bildschirm zur Videoaufnahme: Tasten, dreiteiliger Name, aktuelles Kamerabild und davor laufender Video

Parallel zur Tonaufnahme auf dem iPhone kann das iPad einen Video aufzeichnen. Die Seite für die Videoaufnahme erreicht man mit der „record Video“-Taste auf der iPad-Startseite.

Die Abb. 17 zeigt, was bei der Aufnahme möglich ist.

Man kann die Kamera der Vorder- oder der Rückseite des iPads einschalten. Mit der „record“-Taste nimmt man auf bzw. stoppt die Aufnahme.

Der Video hat einen dreiteiligen Namen bekommen.

Im Moment wurde der Video bereits aufgenommen. Die Datei ist schon in den Cloud-Container hochgeladen worden: „stored!“. „replay“ ist aktiv.

Der gerade aufgenommene Video wird abgespielt (Taste „replay“). Er ist verfügbar, solange er nicht überschrieben wurde.

Im Bild-Hintergrund erscheint der aktuelle Preview der Kamera. Dort würde eine neue Aufnahme ansetzen.

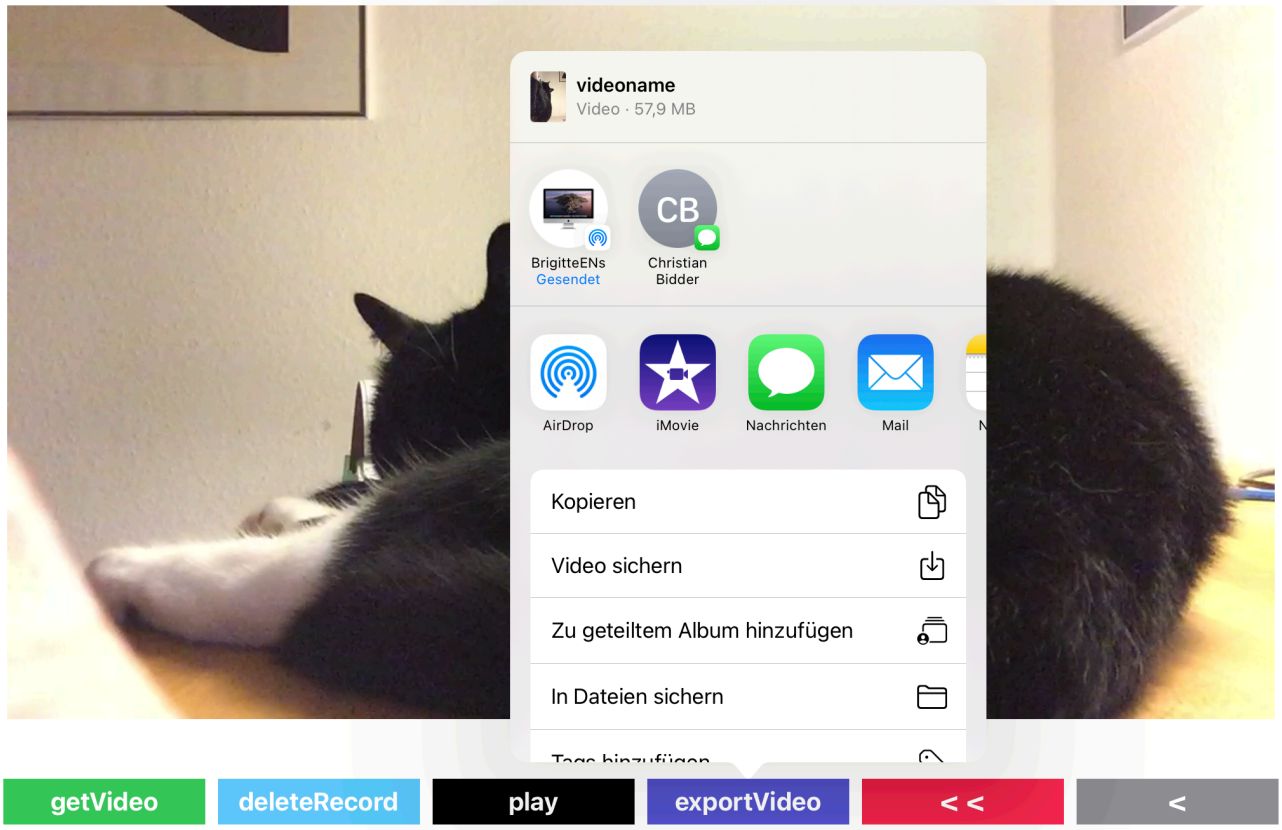

Abb. 18. VideoPlay-Seite: Video holen, abspielen, exportieren und löschen

Die Videos im Cloud-Container werden angezeigt, wenn man in der Übersichtstabelle (s. Abb.16) mit „videoData“ die Liste der Video-Metadaten holt. Tasten rechts führen zur Abspielseite für den gewählten Video.

Abb. 18 zeigt, was man mit dem Video (wie in Abb. 17) dort machen kann.

Dazu holt man die Videodatei zuerst aus dem Cloud-Datensatz ab („getVideo“). Wenn sie angekommen ist, werden die Tasten zu ihrer Verarbeitung aktiv. Man kann den Video abspielen („play“), exportieren („exportVideo“) und löschen („deleteRecord“).

Die Abbildung zeigt, wie der Video exportiert wird, im konkreten Fall auf den iMac der Entwicklerin. Dort ist er bereits angekommen.

Wenn man mit „deleteRecord“ den Video im Cloud-Container löscht, erhält man eine Bestätigung, sobald es passiert ist.

Die Tasten rechts unten führen zurück zur Video-/Audio-Tabelle und zum iPad-Start.

Die Sprachverarbeitung: Übersicht

Die Speech API - die Apple-Spracherkennung - übernimmt die Basis-Transkription der Tondatei. Dabei werden Ergebnisse aus einem breiten Band von Computerlinguistik bzw. Natural Language Processing (NLP) angewandt. Thinkie 3 benutzt deren serverbasierte Version. Sie versorgt mehr Sprachen als die alternative Lösung auf dem Gerät und erreicht eine bessere Qualität.

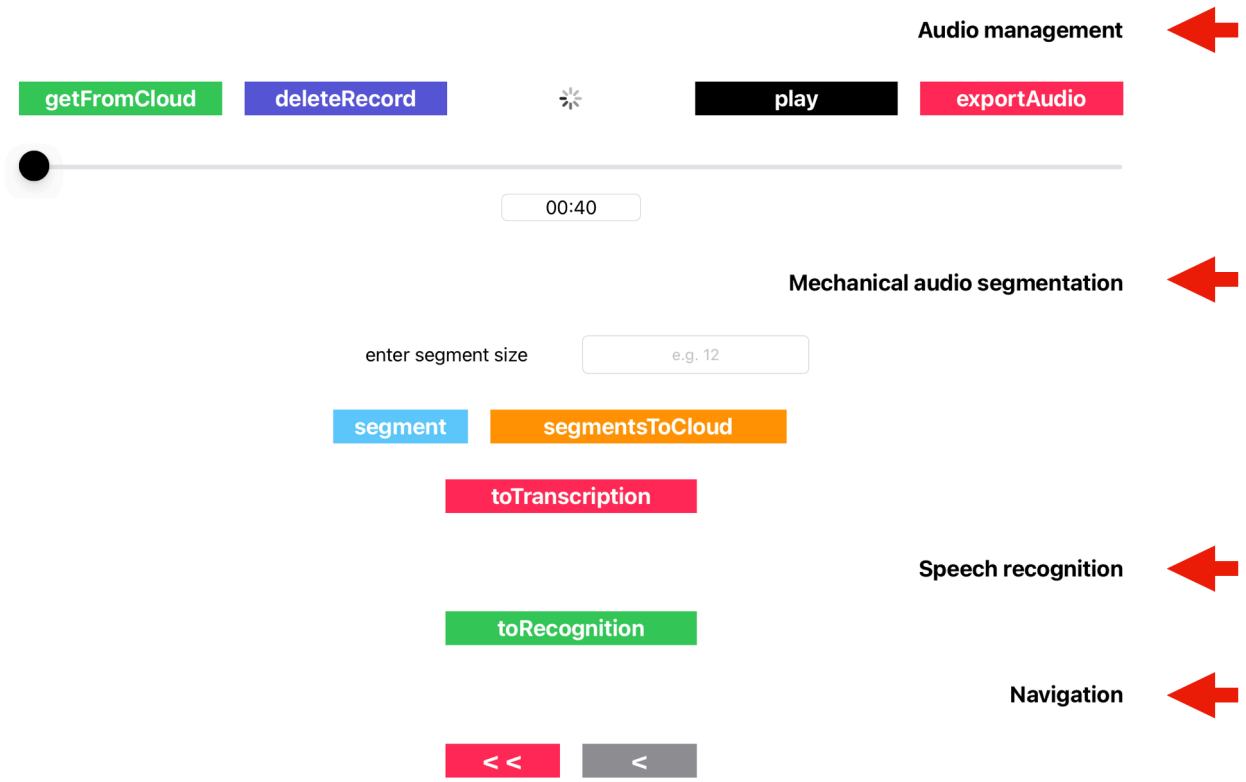

Abb. 19. Speech-To-Text: Optionen auf der Verteilerseite für die Sprachverarbeitung

Selbst wenn man die Anforderungen auf die reine Umsetzung von gesprochener in geschriebene Sprache (Speech-To-Text) einschränkt, muss man mit Schwierigkeiten rechnen. Thinkie versucht es darum mit einem gemischten Einsatz von maschineller Unterstützung und menschlich-intellektueller Transkription und Kontrolle. Transkribiert wird immer in dieselbe Transkript-Datei.

Leistungen, die leicht zu handhaben sind, bietet die Verteilerseite (Abb.19) selbst an:
- die Audiodatei herunterladen, abspielen, exportieren und löschen
- aus der Tondatei Clips zur Bearbeitung ausschneiden. Die Clips der gewünschten Länge kann man im Cloud-Cotainer sichern. Weiter geht es auf der Transkriptionsseite (Abb. 20).
- zur Liste der Audio- und Videodateien und zur iPad-Startseite zurückkehren.

Der Umgang mit der Spracherkennung verlangt von Benutzern mehr eigenen Einsatz. Deswegen ist die gesamte Spracherkennung auf einer eigenen Seite platziert (Abb. 21). Dorthin wird verwiesen.

Wer die Aufgabe kennt, weiß das aus eigener Erfahrung, dass bei Speech-To-Text viel Unsicherheit im Spiel ist. Im Abschnitt über die Transkription findet man mehr dazu.

Die Speech API passt sich an den Sprecher an. Aber auch eine automatische Spracherkennung muss an vielen Stellen interpretieren, wie eine Tonsequenz in den richtigen Text zu transferieren ist. Das funktioniert besser, wenn sie mit Input konfrontiert wird, der Sinn macht.

Thinkie legt es in die Hand der Benutzerin, sinnvolle Tonsequenzen abzugreifen. Die Recognition-Seite unterstützt das Abgreifen der Sequenzen.

Segmentierung

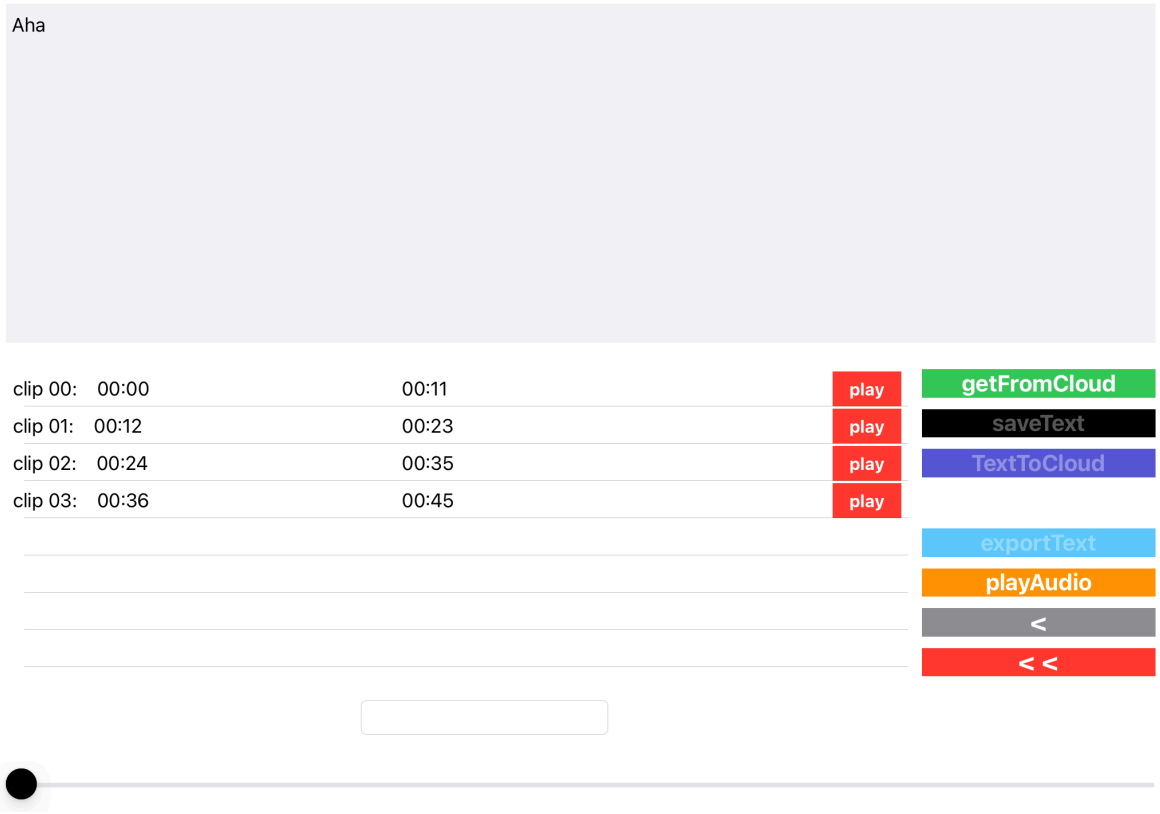

Abb. 20. Segmentierung: Transkript, Liste der Audioclips, Tasten für Audiodatei, Zeitfenster und Regler

Die Segmentierungsseite von Thinkie (Abb.20) stellt oben das Transkript dar. Es ist noch leer.

Mit „getFromCloud“ erscheint dort die Fassung des Transkripts im Cloud-Datensatz. Darunter erhält man eine Liste der Audioclips. Mit ihrem Schalter rechts („play“) spielt man sie einzeln ab. Im Zeitfenster erscheint die erreichte Position.

Man kann das Transkript weiterentwickeln, indem man selbst transkribiert oder indem man den Text kontrolliert und überarbeitet, auch Passagen aus der Spracherkennung.

Das Transkript speichert man lokal mit „saveText“ und mit „TextToCloud“ im Cloud-Container.

Die Taste „playAudio“ spielt die ganze Tondatei ab. Der Gleitregler läuft mit.

Die letzen Schalter führen zurück zur Audio/Video-Liste und zum iPad-Startbildschirm.

Spracherkennung - Speech-To-Text

Für SpeechToText - die Spracherkennung - braucht man sinnvolle Sequenzen aus der Tondatei, die die Erkennung nicht überfordern. Sie sollen maximal 60 Sekunden lang sein. Sinnvolle Einheiten werden besser erkannt.

Erkannt wird die Sprache, die als Systemsprache auf dem iPad eingestellt ist. Wenn man eine andere Sprache will, stellt man die Systemsprache um. Der Verwendung der Spracherkennung muss man zustimmen.

Thinkie greift aus der Tondatei eine angegebene Sequenz ab und schickt sie an den Spracherkennungsserver. Sobald das Ergebnis eingeht, wird es an das Ende des Transkripts angehängt. Dort kann man es so bearbeiten.

Die Teilsequenzen greift man nacheinander beim Abspielen der Datei ab (s. Abb. 21). Mit „play“ beginnt man eine Sequenz, mit „pause“ stoppt man sie.

Die aktuellen Start- und Stoppwerte erscheinen in Start- und Stopp-Feldern auf dem Bildschirm. Wenn man die nächste Teilsequenz abspielt, wird das Ende der Spanne zuvor ins Startfeld geschrieben. Wenn man die Wiedergabe stoppt, wird der Zeitpunkt in das Stoppfeld geschrieben.

Mit „recognize“ schickt man die Spanne von Start bis Stopp zur Spracherkennung. Sobald der Text im Transkript erscheint, kann man zur nächsten Sequenz übergehen.

Nun in Form einer Wiederholschleife:
1. „play“ drücken
2. „stop“ drücken - der Stoppwert erscheint in seinem Feld

Dann für jede Einheit:
3. „play“ drücken - der alte Stoppwert wandert in das Startfeld
4. „stop“ drücken - ein neuer Stoppwert erscheint im Stoppfeld
5. „recognize“ für die Sequenz zwischen Start und Stopp anklicken
6. abwarten, bis das Ergebnis der Spracherkennung an das Transkript angehängt ist, dann eventuell korrigieren
7. zurück zu Schritt 3. : die nächste Sequenz verarbeiten.

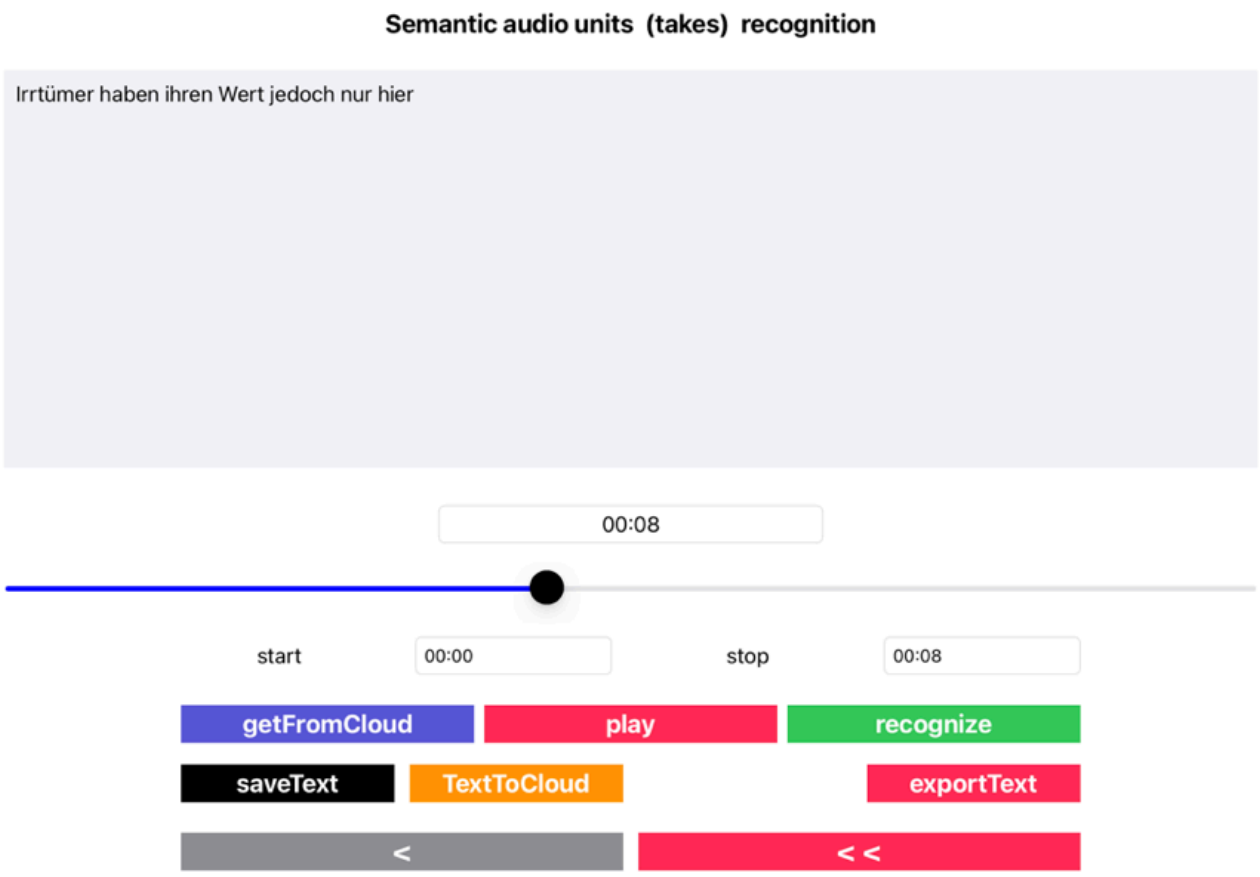

Abb. 21. Speech-To-Text ist im Gange

Zu dem Zeitpunkt, den die Abb. 21 zeigt, hat die Benutzerin bereits einige Schritte erledigt:

- Die Audiodatei wurde aus dem Cloud-Container geholt („getFromCloud“)
- Die Audiodatei wurde kurz angespielt („play“) mit sofortigem Stopp, so dass der Startzeitpunkt für die Spracherkennung auf „00:00“ gesetzt wurde.
- Die Benutzerin hat den Ton laufen lassen, bis eine sinnvolle Einheit abgespielt war und dann gestoppt. Der Stoppzeitpunkt für die Spracherkennung ist mit „08:00“ eingetragen.
- Mit „recognize“ hat die Benutzerin die gewählte Sequenz zum Spracherverarbeitungs-Server geschickt.
- Der Server hat geantwortet. Das Transkript ist in seinem Feld angekommen.

Die Benutzerin geht sie wieder zu Schritt 3 und drückt „play“, um die Datei weiter abzuspielen, maximal 60 Sekunden lang, bis sie die nächste Sinneinheit gehört hat. Dann „stop“ und „recognize“, und so weiter.

Fazit

Lautes Denken

Das Laute Denken von 2021 ist weit über das hinausgewachsen, was die Gründerväter Ericsson & Simon sich 1980 vorgestellt haben.

Lautes Denken Heute hat sich in vielen Fachgebieten verbreitet. Manche von ihnen gab es 1980 noch nicht. Die Verfahren des lauten Denkens haben sich an die Bedingungen unterschiedlicher Disziplinen angepasst.

Im Internet sind für das Laute Denken ganz andere Ressourcen verfügbar als man sich es zu Beginn der Entwicklung vorstellen konnte.

Die Datenaufnahme mit Audio- und Videodatei ist Standard geworden. Eine Blickverfolgung kann hinzukommen. Man kann so vergleichen, was gesagt und was getan wird. Man hat mehr Daten von besserer Qualitäten eine bessere Basis für die Analyse.

Die Denkakte, die aus einem Verbalprotokoll ermittelt werden, haben nach heutigem Verständnis eine innere Struktur, ein eigenes Datenmodell. Datenmodelle passt man an die Merkmale an, die für die Ziele der Untersuchung wichtig sind.

Die arbeitsaufwändige Transkription wird teilweise durch automatische Verfahren der Spracherkennung und der Textanalyse erledigt. Man kann sie an externe Dienstleistungsbetriebe weitergeben.

Verfahren, die bei der Kodierung und Interpretation der Verbalprotokolle helfen, sind verfügbar.

Es gibt erste Ansätze mit Classifiern aus dem Maschinellen Lernen.

Im Vergleich zu 1980 gibt es 2021 Benutzerschnittstellen wie Sand am Meer. Wie deren Benutzer damit zurecht kommen, fragt man aus der Sicht von Benutzern (UX - User Experience).
Lautes Denken ist im Usability Engineering eine der Techniken, mit denen man die Reaktion von Benutzerinnen erfahren kann. Seine Bedeutung steigt.

Was Thinkie 3 zu bieten hat

Thinkie 3 bietet für das Laute Denken eine integrierte mobile Lösung im Handtaschenformat. Sie ist in erster Linie für die Feldforschung gedacht, funktioniert aber überall, wo es ein Netz gibt.

Thinkie läuft weltweit im IOS auf einem iPhone und einem iPad. Die Daten werden in einer Cloud-Datenstruktur gespeichert.

Thinkie 3 konzentriert sich auf die Datenaufnahme mit Audio und Video sowie eine einfache Transkription der Verbalprotokolle mit der Server-Fassung der aktuellen Apple Speech-API.

Alle Daten können exportiert werden. Das ist für eine weitergehende Dateninterpretation erforderlich.

Vergleichbare andere Apps für Lautes Denken gibt es im IOS bisher nicht.

Wie geht es weiter?

Für die Entwicklerin ist es eine ganz praktische Frage, ob sie eine nächste Fassung von Thinkie ins Auge fassen sollte.

Im wissenschaftlich-technische Umfeld herrscht keine Stagnation, also ist dort etwas zu machen. Die Sprachverarbeitung (NLP) und die Anwendungen machen Fortschritte. Die Benutzersicht wird wichtiger, damit auch das Laute Denken.

Verbesserungschancen für Thinkie 4 finden sich leicht: mehr mit der Speech API machen, nach allgemein brauchbaren Merkmalen für die Kodierung / Modellierung suchen, mehr als die Basis-Transkription unterstützen, usw.

Technisch auf der Höhe der IOS-Entwicklung bleiben ist ein Punkt, neue Möglichkeiten zu nutzen ebenso.

Wäre Thinkie mit seinem Ansatz von anderen eingeholt und überholt worden, könnte man ihnen das Feld überlassen. Aber so ist es IOS-weit nicht.

Also mal sehen!!

Literaturangaben

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