ISi SPEECH

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Assistive Technologien in der Sprechtherapie

Juliane Mühlhaus, Hendrike Frieg, Jan Wellmann

17.02.2017

dbs Symposium Bochum

technische universität dortmund

hsg Hochschule für Gesundheit

Fraunhofer IDMT

Was erwartet Sie heute?

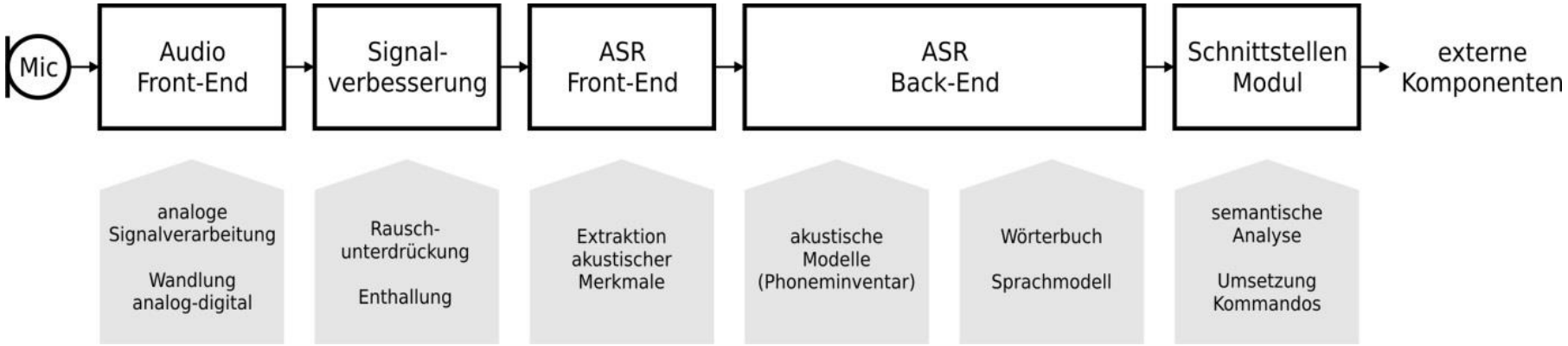
- Einführung in Sprachtechnologien
- Hands-On: Sprachtechnologien
- Apps in der Sprachtherapie
- Hands-On: Einbindung Apps in die Therapie
- Ethisch orientierter Einsatz der Technologien in der Sprachtherapie

Automatische Spracherkennung

Was ist automatische Spracherkennung?

- Der Computer erkennt Sprache durch Vergleich des Eingangssignals mit bekannten Phonemfolgen
  - Erkennerergebnis ist die wahrscheinlichste Phonemfolge
- Mehrstufige Modellierung der Sprache:
  - Erkennung von Lauten (akustisches Modell)
  - Wörtern (Wörterbuch in IPA)
  - Sätze (Sprachmodell)
  - Semantische Analyse (Kognitives Modell, KI)
- Zusätzliche Erfassung von technischen Parametern für den Sprachtherapeuten:
  - Lautstärke
  - Tonhöhe und Melodie (Pitch)
  - Sprechgeschwindigkeit und Pausenlänge

Automatische Spracherkennung



- Beamforming mit bis zu 64 Kanälen
- GSC
- Post-Filter
- …

- MS
- MMSE
- Wiener Filter
- Source Sep. by NMF
- BSS
- …

- GABOR features
- AMS/AMFB
- PNCCs
- MFCCs
- MVN
- RASTA-MFCCs
- FTHs
- …

- Whole-word HMMs
- Triphone HMMs
- Neuronale Netze (MLPs, …)
- SVMs
- Multi-condition training
- MAP and MLLR speaker adaption
- …

- Task Grammar
- *N*-Gram
- Neuronale Netze
- …

Spracherkennung: Wofür?

Interaktion mit Technik

- Auslösen von einfachen Funktionen (‚Licht an‘)
- Computerdialoge (‚Alexa, wie ist das Wetter in Berlin‘)

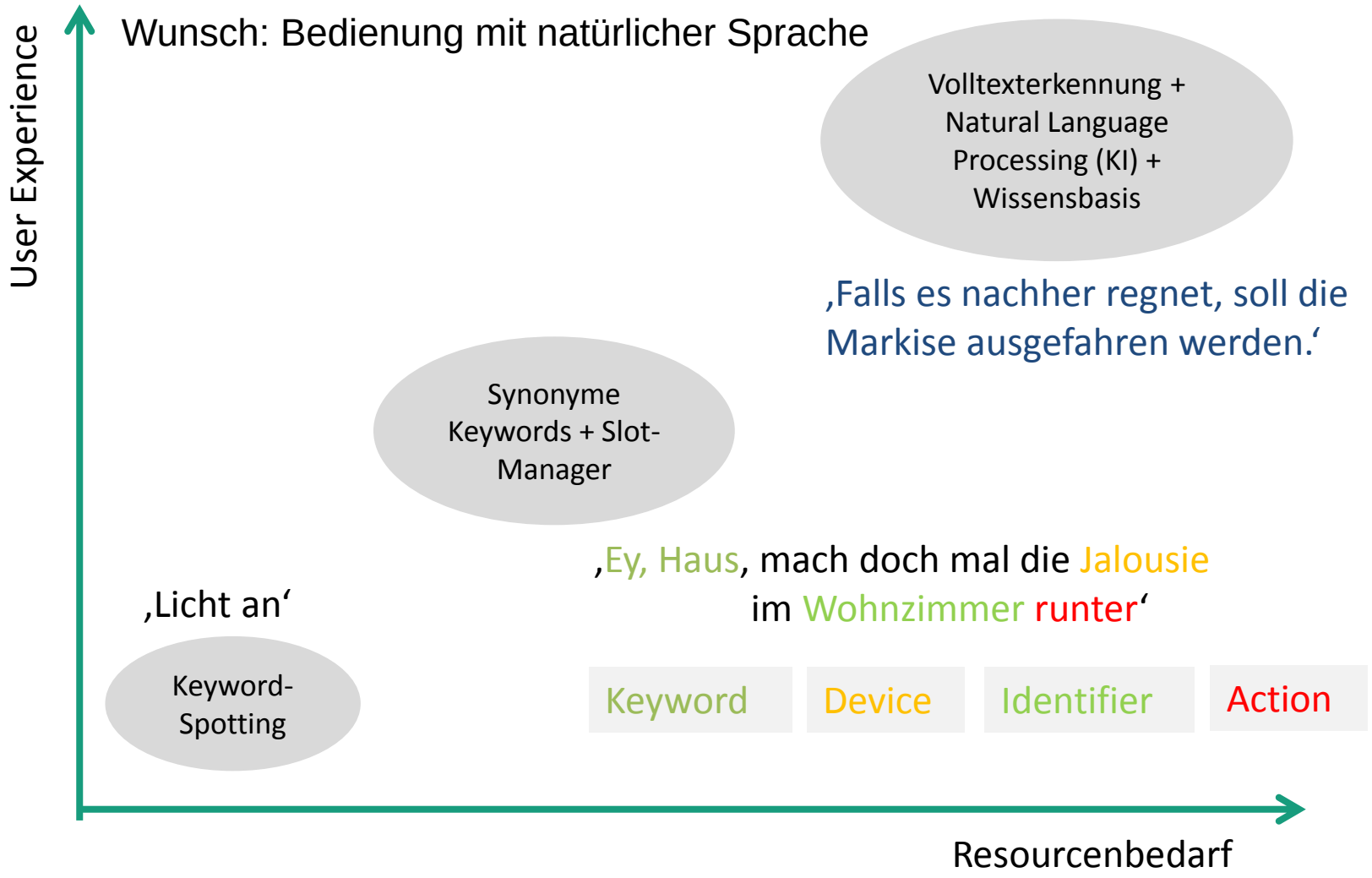
Aussprachetraining

- direkte Nutzung des Erkennerergebnisses zur Bewertung
- weitere sprachtherapeutische Parameter zur Übungssteuerung und zum Feedback

Kommunikationstraining

- Verbales Training von verbaler Kommunkation

Keywords vs. Freie Sprache



Online oder Offline

oder auch Privatsphäre vs. Features

Online

- Skalierbar
- KI mit BigData möglich
- Schwierig: Datenschutz
  - Alle aktuellen Geschäftsmodelle basieren basieren auf Überlassung der Daten
- Leichte Pflege von Modellen und Algorithmen
- Aber auch: Abhängigkeit von externen Services
  - (schlechtes) Beispiel: Revolv

Online oder Offline

oder auch Privatsphäre vs. Features

- Offline
  - kleinere Rechenresourcen (<5W/Node)
  - Begrenzte Wissensbasis (vordefiniert)
  - Maximaler Datenschutz
  - Keine Abhängigkeit von externen Anbietern
  - Updates: schwierig

- Kompromis:
  - lokale WUP-Detektion
  - Cloudbasierte Spracherkennung und Kontexturierung
  - ‚lokale' Cloud: Eigener ASR-Server mit umfangreichen Domänenwissen

‚OK, Google'

Qualitätssicherung

- Qualitätsmaße für Spracherkennung
  - Word Error Rate (WER) oder Word Accuracy Rate (WAR = 100%-WER)

$$WER = \frac{\#INS + \#SUB + \#DEL}{N} * 100\%$$

Referenz		Das	Salz	in	der	Möhren	Creme	Suppe
Erkannt	Ja	Das	Salz		der	Möven	Creme	Puppe
Fehlerart	INS			DEL		SUB		SUB

$$WER=(1+2+1\ )/7*100\%=57{,}1\%$$

- Bei einfacher Kommandoerkenung
  - false positive/negative vs. true positive/negative

Spracherkennung: Probleme

Was kann das Erkennungsergebnis beeinträchtigen?

- Schlechte Signalqualität der Aufnahme:
  - Störgeräusche
  - Mikrofon zu laut / zu leise / defekt
  - Mikrofon ungünstig positioniert

- Sprechbeeinträchtigungen
  - Stark behauchte Aussprache
  - Starker Dialekt
  - Sprechpausen
  - Auslassen von Wortteilen.

Erfordert angepassten Spracherkenner

Technologiegestützte Dysarthrie-therapie: international

- NL: eLearning basierte Sprechtherapie (EST), Beijer et al.
  - Nachsprechen, Abgleich mit Audiofiles
  - Visuelles Feedback zu Intensität und Tonhöhe
  - Testung auditiver Diskriminationsfähigkeit
  - Effekte: Plosivartikulation, Parkinson/PLVT

- EU: Ortho-Logo-Paedia (OLP), Palmer et al.
  - Spracherkennung
  - visuelles Feedback zu Artikulation, Phonation, Tonhöhe und Intensität
  - Effekte: Artikulationspräzision, schwer betroffene Pat.

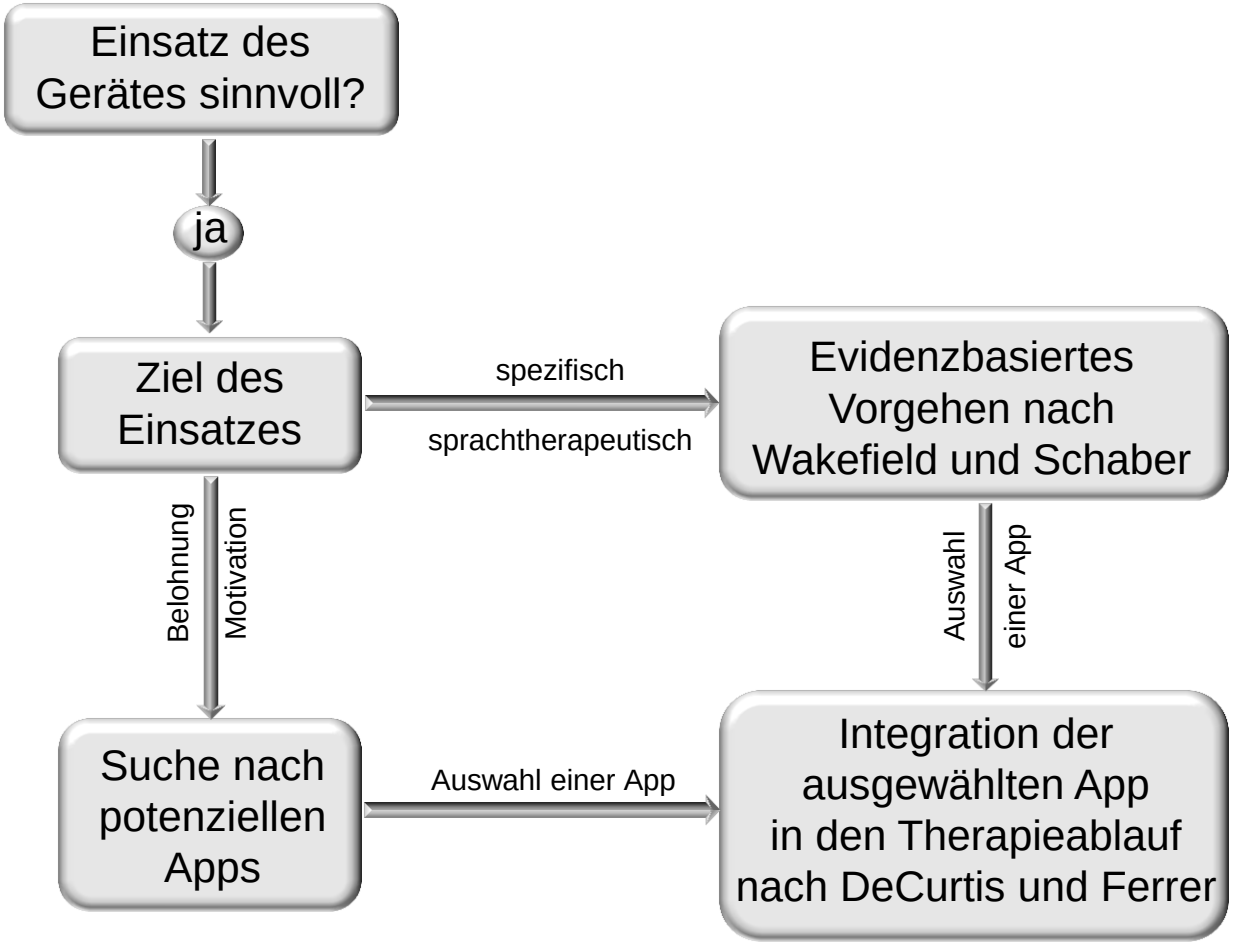
- B: Prosodietrainer, Martens et al.
  - Sprechtempo und Intonation bei Parkinson
  - Auditives Feedback/Wiedergabe eigene Produktion
  - Effekte: Satzverständlichkeit, Pausensetzung, Betonung

Technologiegestützte Dysarthrie-therapie: international

- USA: SpeechVive™, Stathopoulos et al
  - Parkinson, Lombard-Effekt
  - Effekt: Intensität

- S: VoxLog, Gustafsson et al
  - Intensitätsmessung in SPON
  - Taktiles Feedback bei Schwellwertunterschreitung
  - Principles of motor learning
  - Effekte: 3/6dB Unterschreitung nach 500ms
- USA: LSVT® LOUD Companion™, LSVT Global
- USA: EchoWear, Dubey et al
  - Smartwatch, Intensitätsmessung, bisher keine Therapiestudie
- Delayed auditory feedback
  - Keine Transfereffekte

Richtlinien zur Auswahl von Apps



(Starke & Mühlhaus, 2016)

Evidenzbasiertes Vorgehen nach Wakefield & Schaber

- **Schritt 1: Formulieren der klinischen Frage**
  - Ausgangspunkt für Auswahl einer App ist Frage
  - Konkretisierung der Frage durch klinische Fragestellung (PICO)

(Wakefield & Schaber, 2012)

Evidenzbasiertes Vorgehen nach Wakefield & Schaber

- **Schritt 1: Formulieren der klinischen Frage**
- **Schritt 2: Evidenz finden**
  - Recherche in einschlägigen, frei zugänglichen Datenbanken oder Plattformen für evidenzbasierte Praxis
  - z.B. SpeechBITE, evidenssst.org, Cochrane Library, ASHA, SpeechandLanguage.com, Google Scholar

(Wakefield & Schaber, 2012)

Evidenzbasiertes Vorgehen nach Wakefield & Schaber

- **Schritt 1: Formulieren der klinischen Frage**
- **Schritt 2: Evidenz finden**
- **Schritt 3: Bewertung der Evidenz**
  - Bewertung der identifizierten Artikel o.ä. Quellen hinsichtlich ihrer Passung zur Fragestellung, Güte und Evidenzstufe

(Wakefield & Schaber, 2012)

Evidenzbasiertes Vorgehen nach Wakefield & Schaber

- **Schritt 1: Formulieren der klinischen Frage**
- **Schritt 2: Evidenz finden**
- **Schritt 3: Bewertung der Evidenz**
- **Schritt 4: Suche im App-/GooglePlaystore und Abgleich mit der gefundenen Evidenz**
  - Suche nach potenziellen Apps im jeweiligen Store des genutzten Gerätes
  - Bewertung der gefundenen Apps durch Vgl. App-Beschreibungen mit gefundener Evidenz aus Schritt 4
  - Übereinstimmung zwischen Inhalt und Format der App mit Aspekten der evidenzbasierten Intervention (Schritt 4)?
  - Übereinstimmung Zielgruppe der App mit Patient*in?
  - Ist Kostenaufwand gerechtfertigt?

(Wakefield & Schaber, 2012)

Evidenzbasiertes Vorgehen nach Wakefield & Schaber

- **Schritt 1: Formulieren der klinischen Frage**
- **Schritt 2: Evidenz finden**
- **Schritt 3: Bewertung der Evidenz**
- **Schritt 4: Suche im App-/GooglePlaystore und Abgleich mit der gefundenen Evidenz**
- **Schritt 5: Klinische Entscheidung treffen**
  - Kann eine App in die Therapie integriert werden?
  - Wenn ja, welche App kann in die Therapie integriert werden?

(Wakefield & Schaber, 2012)

Therapieablauf nach DeCurtis & Ferrer

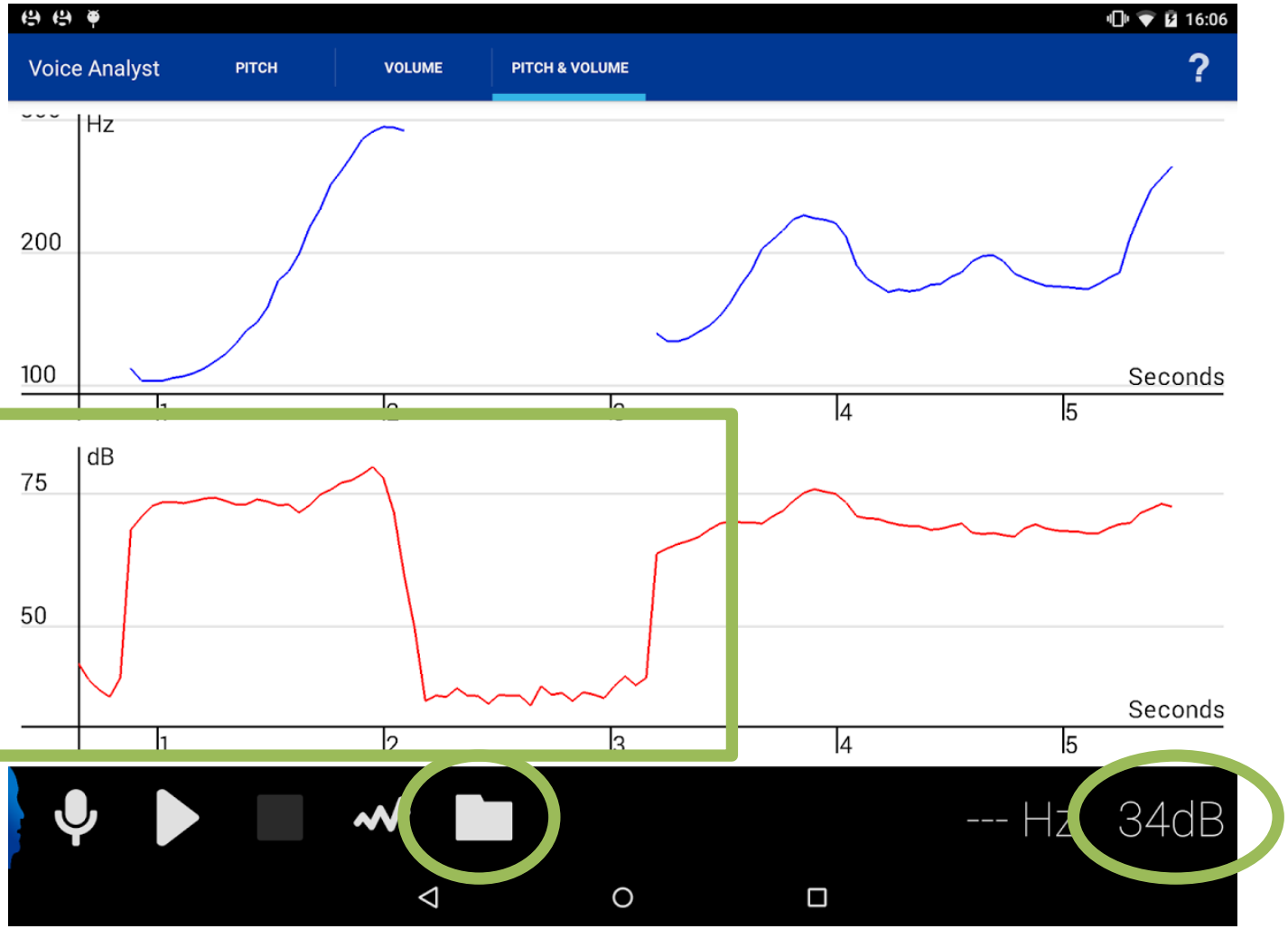
- **Vorbereitung:** Welche Gründe gibt es für die Integration eines mobilen Gerätes gegenüber traditionellem Therapiematerial in der Therapie?
- **Zielgruppe:** Alter, Geschlecht, Beeinträchtigung, Schweregrad der Beeinträchtigung der/des Patient*in? Soll das Gerät in Einzel- oder Gruppenkontexten verwendet werden?
- **Zeit:** Wie viel Zeit soll der/die Patient*in maximal mit dem Gerät üben?
- **Zweck:** Für welchen Zweck wurde die App konzipiert? Inwieweit passt diese mit den individuellen Zielen der/des Patient*in zusammen?
- **Positionierung:** Welche Effekte gibt es, wenn Patient*in und Therapeut*in nebeneinander sitzen im Vergleich dazu, wenn sie gegenüber sitzen?
- **Potenzial:** Wie wird der Therapiefortschritt mit der App in alltägliche Kontexte übertragen? Wo gibt es Bezugspunkte im bisherigen Wissens des/der Patient*in und dem Therapieinhalt?

(nach DeCurtis & Ferrer, 2015)

Apps in der Dysarthrietherapie

Android

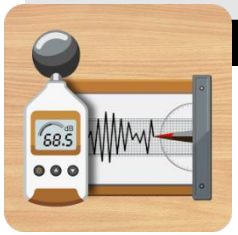
- Voice analyst (4,99€/kostenfrei), Speech Tool Limited, Bristol/UK

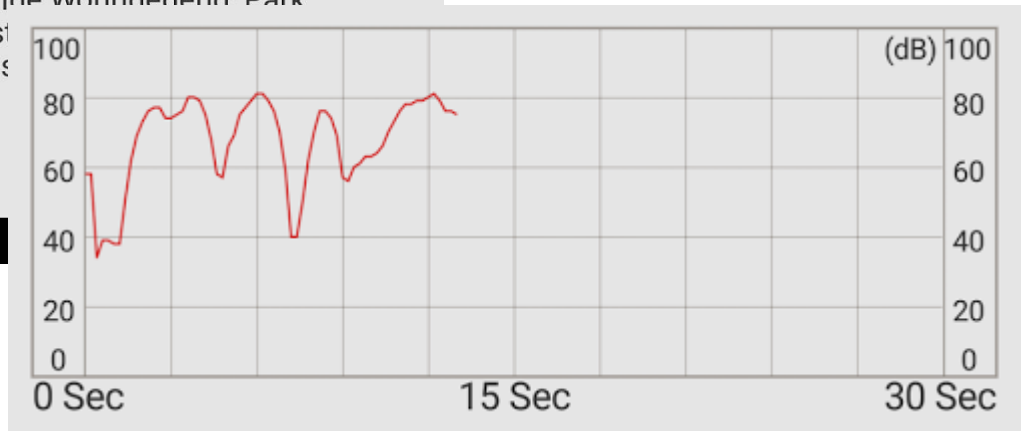


Apps in der Dysarthrietherapie

Android

- Schallmessung: Sound Meter Pro (1,50€/kostenfrei), Smart Tools Co., Korea

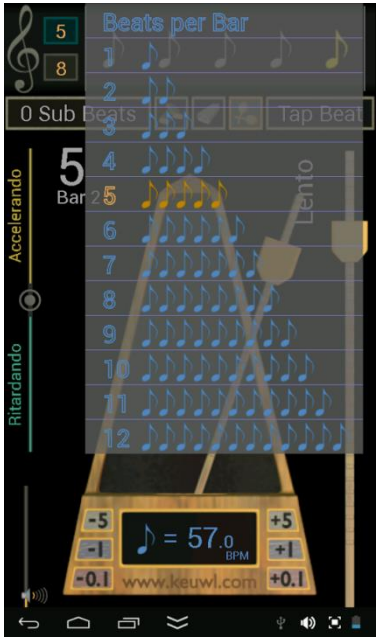


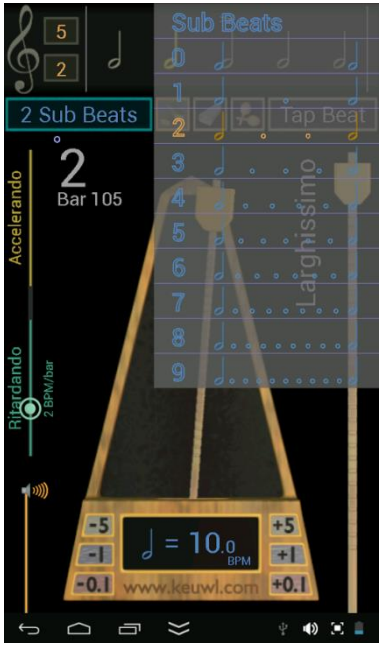


Apps in der Dysarthrietherapie

Android

- Metronom
  - z.B. keuwlsoft, kostenfrei

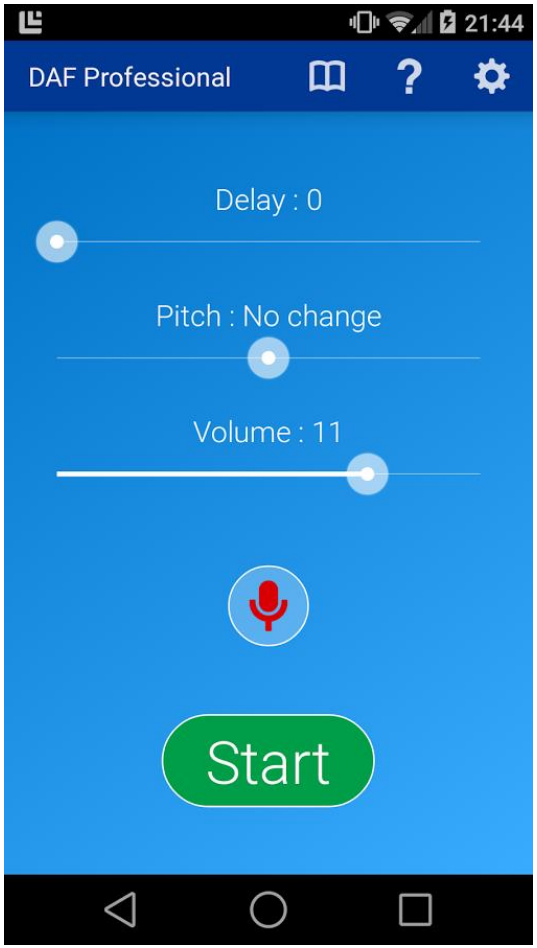




Apps in der Dysarthrietherapie

Android

- DAF (delayed auditory feedback) (3,31€/kostenfrei), Speech Tool Limited, Bristol/UK

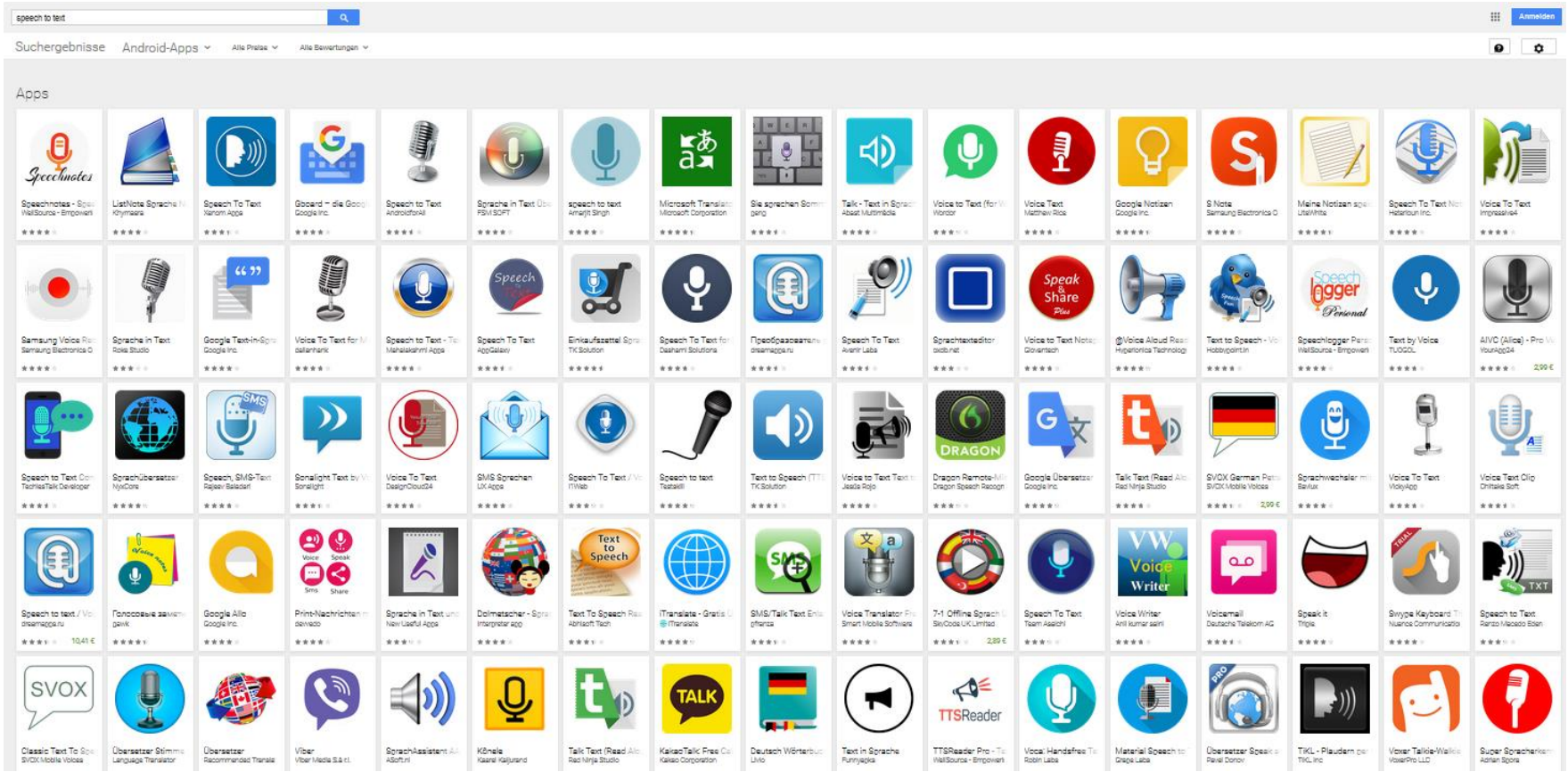


Apps in der Dysarthrietherapie

Android

- Diktierfunktion des Smartphones zur Aufnahme/Wiederabspielen

- Speech-to-Text/Spracherkennung …



Ethisch orientierter Einsatz der Technologien in der Sprachtherapie

- MEESTAR

  (**M**odell zur **e**thischen **E**valuation **s**ozio**t**echnischer **Ar**rangements)

Abbildung 1:
MEESTAR: x-Achse: Dimensionen der ethischen Bewertung; y-Achse: Stufen der ethischen Bewertung; z-Achse: Ebenen der ethischen Bewertung.

(aus: Manzeschke et al., 2013)

Ethisch orientierter Einsatz der Technologien in der Sprachtherapie

ISi SPEECH

- **Leitsätze**

1. Der Einsatz Neuer Technologien in der Sprachtherapie soll dazu beitragen, dass jede/r Nutzer*in ein selbstbestimmtes Leben führen kann.
2. Neue Technologien sollen bei kognitiv beeinträchtigten Personen in der Sprachtherapie nur eingesetzt werden, wenn eine gesonderte Prüfung stattgefunden hat und der (mutmaßliche) Wille der Person berücksichtigt wird.
3. Der Einsatz Neuer Technologien in der Sprachtherapie soll die Teilhabe am gesellschaftlichen Leben und die Integration in die gesellschaftlichen und sozialen Verbindungen unterstützen.
4. Der Zugang zu Neuen Technologien für den Einsatz in der Sprachtherapie soll diskriminierungsfrei gestaltet werden.
5. Der Umgang mit den Neuen Technologien muss für alle Nutzerinnengruppen sicher und frei von Belastung sein. Dies bezieht sich auf die normale Anwendung, auf potenzielle Fehler und Ausfälle der gesamten Technik sowie auf einzelne Prozessketten.
6. Die erhobenen Daten beim Einsatz Neuer Technologien in der Sprachtherapie sind durch besondere Schutzmaßnahmen abzusichern.

(Mühlhaus & Hastall, 2016; modifiziert nach Manzeschke et al., 2013)

Ethisch orientierter Einsatz der Technologien in der Sprachtherapie

- **Leitsätze**

7. Nutzerinnen sollen über Datenerhebung und Funktion der Neuen Technologie aufgeklärt und informiert sein, bevor sie eine Entscheidung über den Einsatz der Neuen Technologien in der Sprachtherapie treffen.
8. Verantwortlichkeiten und Zuständigkeiten zur Neuen Technologie und seiner Funktion sind klar definiert.
9. Stigmatisierung und Diskriminierung sind beim Einsatz Neuer Technologien ebenso unerwünscht wie von der Technologie angenommene Normierungen.
10. Der Einsatz der Neuen Technologien ist für die Nutzerin einfach, intuitiv und nachvollziehbar.
11. Die Nutzung der Neuen Technologie soll jederzeit beendet werden können.
12. Alle Beteiligten im Bereich der Neuen Technologien sollen entsprechend qualifiziert und weitergebildet sein.
13. Der Nutzen und Mehrwert Neuer Technologien ist für alle Beteiligten erkennbar und nachvollziehbar

(Mühlhaus & Hastall, 2016; modifiziert nach Manzeschke et al., 2013)

Kontaktdaten

- Hendrike.fried@hs-gesundheit.de
- Juliane.muehlhaus@tu-dortmund.de
- Jan.wellmann@idmt.fraunhofer.de