ANWENDUNGSORIENTIERTER EINSATZ VON SPRACHERKENNUNG UND SIGNALVERARBEITUNG
KONKRETE PROJEKTE, PROJEKTSKIZZEN UND IDEEN

Stieger Franz,

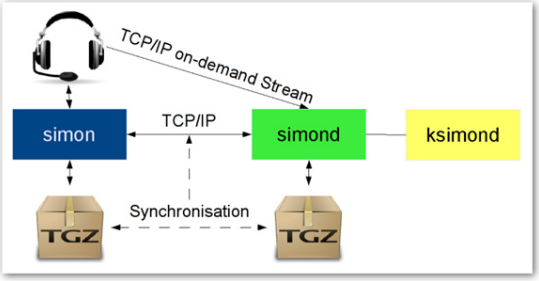
Simon Listens gemeinnütziger Verein für Forschung und Lehre EU-PIC: 987058033
www.simon-listens.org www.simon-listens.com
f.stieger@simon-listens.com

Abstract: Grundsätzlich ist es für mich immer irritierend zu Konferenzen über „Artificial Intelligence“ oder „Sprachsignalverarbeitung“ als Sprecher eingeladen zu werden, da meine ursprüngliche Tätigkeit als Sonderschullehrer und Sprachtherapeut, dies nicht vermuten lassen würde. Die hauptsächliche Begründung dieser Einladungen sehe ich daher in meinen Erfahrungen und meiner Beteiligung im Bereich der Projektideen, Projektentwicklungen und meiner anwendungsorientierten Sichtweise vieler technischer Lösungen. Der Prozess der Analyse ist im Bereich der Forschung vor allem in Europa sehr dominant, und der Bereich der Synthese, welche intensive interdisziplinäre Zusammenarbeit und Kommunikation unter den verschiedensten Forschungsinstitutionen erfordern würde, ist sehr unterentwickelt. Aus diesen Gründen werde ich hier in diesem Beitrag in erster Linie konkrete Entwicklungen, Anwendungsfälle, Lösungen, Forschungsprojekte, Produktentwicklungen und noch nicht umgesetzte Projektskizzen und Fragestellungen für die Zukunft beschreiben, um so eventuelle Impulse und Anregungen für mögliche Projekte der Zukunft setzen zu können. Der Ursprung aller dieser Projekte ist in meiner sprachtherapeutischen Arbeit in Kombination mit meinen IT-Kenntnissen und Erfahrungen zu suchen und bestand in der ursprünglichen naiven Illusion, für körperbehinderte Schüler und Personen Spracherkennung nutzbar zu machen, da in diesen Fällen der Bedarf am größten ist. Diese Zielgruppe wurde durch den Bereich der Senioren erweitert, da die Situation pflegebedürftiger Senioren oftmals und praktisch dieselbe ist, und da viele Studien aussagen, dass neben den physischen und funktionellen Defekten, die Sprache am natürlichsten ist und auch am längsten erhalten bleibt.

1 Softwareentwicklung und Forschungsprojekte

1.1 Simon Open Source Speech Recognition (www.simon-listens.org)

simon ist das Haupt-Frontend zu der simon Open-Source Spracherkennungslösung. Es ist ein simond Client und bietet ein grafisches Interface um das Sprachmodell und die Kommandos zu verwalten.



1.2 Simon Tools

1.2.1 SAM – Simon Acoustic Modeller

Sam ist eine Software zum Erstellen und Testen akustischer Sprachmodelle. Es kann neue Sprachmodelle kompilieren, von Simon erstellte Modelle benutzen und selbst Modelle für Simon kreieren.

1.2.2 SSC – Simon Sample Collector

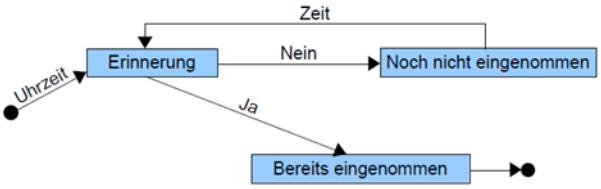
SSC ist eine Software für Sprachaufnahmen in großem Ausmaß. Mit SSC können mehrere Teams Trainingsdaten von potentiellen Endnutzern oder professionellen Sprechern aufnehmen und diese am zentralen SSCD Server speichern.

1.2.3 Afaras

Afaras ist ein Tool mit dem fehlerhafte oder qualitativ minderwertige Aufnahmen aus Datensets, die mit SSC erstellt worden sind, effizient entfernt werden können.

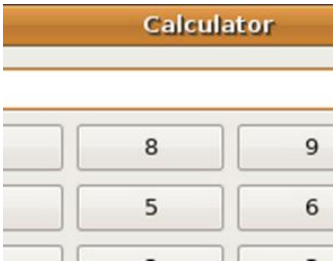
1.2.4 Simon Dialogsystem

Im Rahmen unserer Forschungsprojekte wie Benefit und Astromobile wurde ein bidirektionales Dialogsystem entwickelt, welches in erster Linie für Betreuungsinstitutionen und Angehörige, die Erinnerungsfunktionen und einfache Dialoge für die Betreuung und Interaktion nutzen wollen, von Interesse ist.



1.2.5 Simon Basisautonomie – Versuch von einfachen Lösungen für körperbehinderte Personen

Neben der Steuerung der Maus über ein Desktopgitter sind folgende Funktionen individuell einstellbar. Mit wenigen frei wählbaren Wörtern aus der Alltagssprache wie „rechts“, „links“, „rauf“, „runter“, „Ok“, „stop“ etc. ist es möglich seine eigene Musik, Diashows, Videos, Fernseh- und Radiosender etc. anzusteuern. Das tägliche Lesen der Zeitungen und das Anwählen anderer Internetadressen ist einfach und unkompliziert möglich. Mit dem Zahlen-Plug-In in Firefox kann durch die Eingabe von Zahlen jeder Link einer Webseite angewählt werden. Mittels Zahlen können vorgegebene E-Mail-Adressen angewählt, unter Verwendung von frei erweiterbaren Textbausteinen die wesentlichen Fragen gestellt und Informationen über sich selbst versendet werden. Unter Verwendung von Skype oder einer anderen Voice Over-IP Lösung ist es einfach Telefonverbindungen mit Verwandten und Freunden in der ganzen Welt herzustellen. Ergänzt wird dieses Anwendungspaket noch mit einem verbal steuerbaren Rechner, der sowohl für Rechenoperationen als auch für den pädagogischen Einsatz, d.h. der Möglichkeit die Rechenoperationen darzustellen, einzufügen und auszudrucken, geeignet ist.



1.2.6 Simon Szenarien

Da die Anwendungsfälle für jede Person sich ständig verändern, war ursprünglich daran gedacht, Szenarien zu entwickeln, die online für die Verwendung verfügbar sein sollen. Mit Szenarios lassen sich gewisse Anwendungsfälle in Simon gruppieren und diese „Bausteine“ kombinieren. Damit lassen sich allgemeine Pakete für individuelle Lösungen zusammenstellen und anpassen. Mangels Ressourcen konnte dieser Ansatz nicht weiter verfolgt werden. Es ist jedoch daran gedacht in einem zukünftigen Projekt Szenarien für die Steuerung von Lernsoftware für körperbehinderte Schüler online zur Verfügung zu stellen.

1.3 Simon Forschungsprojekte

1.3.1 Projekt Benefit: Verbale Steuerung von IKT Anwendungen von Senioren

Dies war das erste Forschungsprojekt mit Senioren und der Steuerung von Szenarien der Information und Kommunikation über den Fernseher. Die Ergebnisse dieser Arbeit sind in dem Video https://www.youtube.com/watch?v=5h-gSADR8mg und dem Sachdokument http://www.simon-listens.org/index.php?id=407 im Detail ersichtlich und frei verfügbar. Teile dieser Ergebnisse sind in das Projekt Simon Touch eingeflossen.

1.3.2 EU-Projekt ASTROMOBILE im Rahmen des ECHORD Programmes

Das Projekt ASTROMOBILE mit der Entwicklung von Simon Touch, der Sprach- und Dialogsteuerung des Roboters ASTRO in Deutsch Englisch und Italienisch, der Entwicklung von Owncloud Tools für die Planung von Terminen, Dialogen, der Entwicklung eines Meego Clients für das Rufen des Roboters und webbasierte Kontrolltools für die Steuerung des Roboters ASTRO hat uns Europaweit einen großen Imagegewinn gebracht. Die Ergebnisse dieser Arbeit sind in dem Video https://www.youtube.com/watch?v=PDkEe4G4PX4 und dem Sachdokument http://www.simon-listens.org/index.php?id=406 im Detail ersichtlich und frei verfügbar. Teile dieser Ergebnisse sind in das Projekt Simon Touch eingeflossen.

1.3.3 Projekt Benefit: „ ibi“ ich bin informiert

Das AAL-System „ibi“ soll für Menschen mit Pflegebedarf und deren Angehörige sowie PflegerInnen (mobiler Dienst) als innovatives Informations- und Kommunikationsmittel dienen. Ein für TV & Tablet optimiertes, sprachgesteuertes animiertes Avatar- Interface erleichtert die Kommunikation (z.B. SMS) und stellt Information wie Pflegehilfestellungen, Kalender und Open Data Infos attraktiv und benutzerfreundlich dar. Genauere Informationen sind auf der Webseite http://www.ibi.or.at/ verfügbar

2 Businesslösungen und Produktentwicklungen

2.1 Simon Touch mit Owncloud Verwaltung und Simon Micro

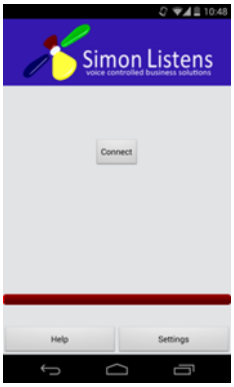
SimonTouch ist eine verbal steuerbare Kommunikationsstation, welche unter Verwendung natürlicher Wörter wie rauf, runter, ok, rechts, links etc. bzw. mittels Touchscreen bei Vorhandensein eines touch-fähigen Monitors auf jedem handelsüblichen Fernseher einsetzbar ist. Simon Touch ist in einer virtuellen Box konfiguriert, sodass diese nur auf einen Rechner kopiert, die Internetverbindung und die Owncloudverbindung konfiguriert werden muß, um die Funktionalität und Verwaltung sicherzustellen. SimonTouch umfasst dabei folgende Funktionen:

- Hilfe/Telefonkontakt zu Hausarzt, Rettung, Pflegeperson und Vertrauensperson
- Familienfotos
- Familienvideos
- Eigene Musik
- Nachrichten wesentlicher Zeitungen
- Telefon, Skype Gespräche und e-Mailempfang (nur lesen – schreiben bei Touchscreen)

- Individueller Terminplan und Planung der Erinnerungsdialoge
- Herstellung von Telefonverbindungen für Einkauf (Kaufmann/Apotheke) und Transport (Taxi, Rettung und persönlicher Transport)

Die OwnCloud-Verwaltung steht als Backend für jegliche Kommunikations- und Informationsstationen zur Verfügung. Der Austausch von Multimediadateien (Bilder, Musik, Videos) und das Erstellen neuer Termine, sowie das Verwalten der Skype-Kontakte muss nun also nicht mehr bei der lokalen SimonTouch-Station selbst vorgenommen werden sondern erfolgt online.

SimonMicro (Im Playstore erhältlich) ist eine Android Applikation, die es dem Patienten ermöglicht, Android Geräte, wie z.B Smartphones und Tablets, als drahtloses Mikrofon zu verwenden. SimonMicro bietet zwei unterschiedliche Funktionsweisen, wie die Sprachaufnahmen erfolgen können:



Push-Funktion:

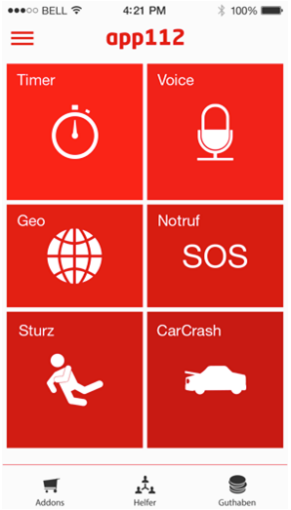
Bei der Push-Funktion muss man, um die Sprachaufnahme zu starten, einen Button gedrückt halten. Erst beim Loslassen wird die Aufnahme an Simon geschickt und zur Spracherkennung weiterverarbeitet.

Activate-Funktion:

Für Personen mit eingeschränkter Motorik gibt es die Activate-Funktion. Hierbei ist die Sprachaufnahme nach einmaliger Aktivierung ständig aktiv und sendet die Aufnahmen an Simon, wenn ein gewisser Lautstärkepegel überschritten wurde.

Diese Entwicklungen sind zum Großteil verfügbar, aber aufgrund unserer minimalen Organisationsstruktur, fehlender Kontakte und geringen finanziellen Möglichkeiten nur schwer zu vermarkten.

2.2 App112 und darauf aufbauendes Produktportfolio



App112 ist eine Alarm-App mit der man auf 8 unterschiedliche Arten einen Alarm auslösen kann. Mit App112 wird zunächst das enge, soziale Umfeld einer Person angesprochen. Familienmitglieder, Freunde oder Nachbarn erhalten sofort eine Benachrichtigung per SMS, E-Mail oder Telefon, sobald ein Alarm durch App112 ausgelöst wird. Dadurch sind die Alarm auslösende Person und die entsprechenden Helfer - die Alarmempfänger - in der Lage, sofort in Kontakt zu treten und mittels Geotracking ebenfalls die Lokalisierung des Hilferufenden genauestens festzustellen. Die Plattformunabhängige Programmierung ermöglicht es die App auf verschiedensten Geräten und Plattformen zu betreiben. Die Smartphone Versionen sind im Playstore für Android und Apple Appstore verfügbar. (http://app112.com/)

2.2.1 Signalverarbeitunglösungen im Rahmen des App112

Im Rahmen der Algorithmenentwicklung wurden für das Projekt App112 3 verschiedene akustische Alarmkanäle entwickelt:

a. Kanal für Schlüsselworterkennung
b. Kanal für Rhythmuserkennung
c. Kanal für Alarmmeldererkennung

Jeder Kanal ist mit unterschiedlichen Hintergrundgeräuschklassifizierungen ausgestattet, um abhängig von der Hintergrundgeräuschsituation die Berechnungsmodelle anzupassen, und/oder teilweise auch nur um Rechenpower zu sparen. Für den Referenzaufnahmeprozess selbst wurden zusätzliche Qualitätskriterien und Rückgabewerte für den User implementiert, um die Aufnahme der Referenzen optimal zu gestalten. Die große Herausforderung in diesem Projekt der Produktentwicklung ist die Auflösung des Spannungsfeldes zwischen Optimalem Erkennungsergebnis in jeder Situation und geforderten minimalen Fehlalarmen für jeden Kanal, und dies vor dem Hintergrund, dass die App auf vielen verschiedenen Smartphones mit unterschiedlichster Soft- und Hardwareausstattung optimal funktionieren sollte. Für die Umsetzung dieses Projektes wurde im Entwicklungsbereich intensiv mit dem SPSC-Lab der Technischen Universität Graz zusammengearbeitet und ebenfalls eine Diplomarbeit für die Entwicklung des Keywordspotting Algorithmus und der Hintergrundklassifizierung umgesetzt. Siehe Literaturliste!

2.2.2 Segfi Button und Base

Wer kein Smartphone besitzt oder eine Alternative für Zuhause möchte, kann sich die kompatible app112-Base ganz einfach in eine Steckdose in den eigenen 4 Wänden stecken. Die app112-Base benötigt lediglich eine zuverlässige Stromversorgung und übernimmt die Funktion ihres Smartphones. Der Fokus der app112-Base liegt dabei auf dem Erkennen von "Hilferufen", Klopfzeichen oder dem Drücken des SOS-Knopfes. Es können bequem 3 verschiedene Alarmkanäle mit Referenzwörtern oder Klatschrhythmen belegt werden. SegFi ist ein zusätzlicher externer Alarm Knopf, der mit der app112-Base kompatibel ist. Dank Knopfdruck und die dadurch umgehende Verbindung zur app112-Base, lassen sich Alarme einfach und schnell auslösen.

Verteilen Sie SegFi-Knöpfe in Ihrer Wohnung oder im Büro – so sind Sie jederzeit sicher und in der Lage, ihre Helfer im Notfall zu benachrichtigen. (http://www.segfi.com/)

2.2.3 Fireflag

Fireflag ist eine etwa kartenspielgroße Box, die an jede Steckdose angeschlossen werden kann und dann die umliegenden Räume akustisch überwacht. Sobald in einem der überwachten Räume ein akustischer Rauchmelder anschlägt, erkennt Fireflag das Rauchmeldersignal und benachrichtigt telefonisch und/oder per SMS unmittelbar bis zu 5 vorher festgelegte Alarm-Empfänger (z.B. das Mobiltelefon der Eigentümer/Mieter, die Nachbarn, den Immobilienverwalter, einen Sicherheitsdienstleister, etc.). Fireflag kann Sie außerdem nicht nur vor Feuer schützen. Das Gerät kann neben dem Rauchmeldersignal noch zwei weitere Signale speichern. So können Sie beispielsweise noch ein Glasbruchsignal für den Fall eines Einbruchs und ein Wassermeldersignal bei Wasserschäden auf dem Gerät aufnehmen. (http://www.fire-flag.com/)

2.2.4 Vitalshirt

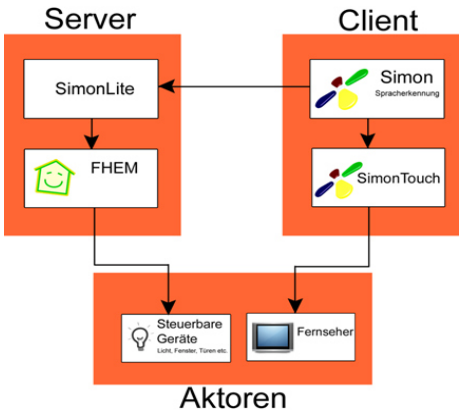
Zusätzlich zu den komplementären Angeboten werden noch Vitalshirts entwickelt, die mit App112 in Verbindung stehen und verschieden biometrische Daten übermitteln können wie z.B. Herzfrequenz-Atemsensor-Kombination, die erkennt u.a. Herz-Unregelmäßigkeiten, Panik, Herzinfarkt. Sturzsensor etc.

3 Neue Projektideen und Projektskizzen

3.1 Verbal steuerbares Kranken/Pflegezimmer

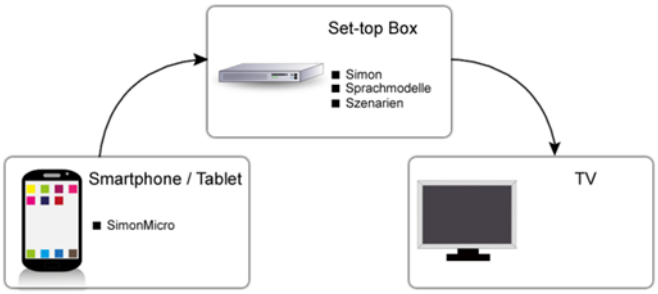
Die wesentlichen Schwerpunkte des Projektes „Verbal steuerbares Kranken-Pflegezimmer“ beschäftigen sich damit, mit Hilfe der Spracherkennungssoftware Simon und Simon Touch, neben den bereits bestehenden Informations- und Kommunikationsmöglichkeiten (Fotos, Videos, News, Skype und Telefonmöglichkeit etc.) ebenfalls die verbale Steuerung von Aktoren in einem Kranken-Pflegezimmer bzw. privaten Wohnbereich ausführen zu können. Der Kunde soll dazu in der Lage sein mit Hilfe der entwickelten Spracherkennungssoftware und Simon Touch folgende Aktoren steuern zu können. Der Schwerpunkt liegt dabei auf das Bedienen folgender Geräte:

- Licht
- Eingangstür
- Tür
- Fenster
- Jalousien
- Krankenbett
- Fernseher
- Schwesternruf



3.2 Verbal steuerbare Set-Top Box

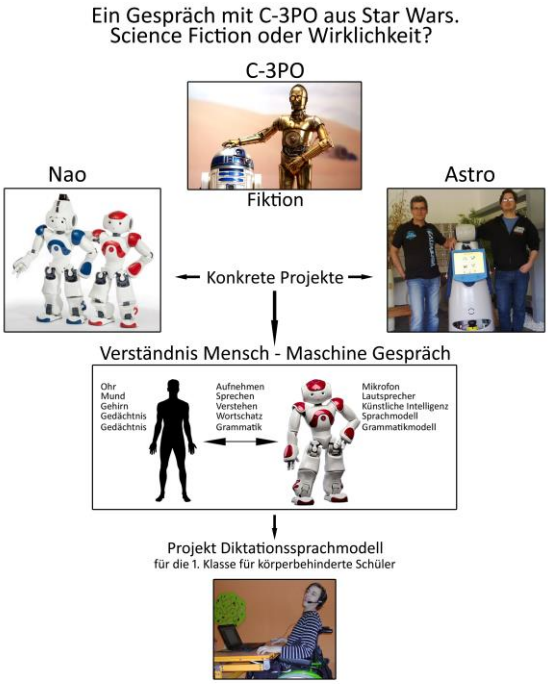
„Simon – verbal steuerbare Set-Top Box“ ist eine Projektidee zur generellen Lösung für die erleichterte Bedienung des Fernsehers für Senioren mit Seh- und Feinmotorischen Problemen, und motorisch behinderte Personen. In Zukunft soll aus Rücksicht auf die enorm große Kundenstruktur der SeniorInnen, optional für körperbehinderte Personen, und zur Gewinnung von Marktvorteilen Set-Top Boxen mit Spracherkennung aufgerüstet werden, um über diese Schnittstelle den Senioren und körperbehinderten Personen möglichst lange eine optimale Steuerung der Fernsehbedienung zu ermöglichen.



Natürlich wäre es in Zukunft ebenfalls möglich Fernsehgeräte an sich verbal steuerbar zu machen, da ja bereits viele Geräte über Internetzugang und genügend Computerressourcen verfügen.

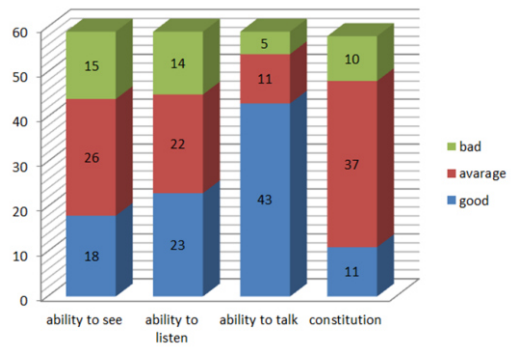
3.3 Talente Regional – Diktationsmodell für körperbehinderte Schüler

Einsatz der Open Source Spracherkennungssoftware Simon für die Steuerung von Lernsoftware und Entwicklung einfacher Diktationsmodelle basierend auf dem Wortschatz und Grammatik in der 1. und 2. Schulstufe unter aktiver Forschungsteilnahme von Schülern der genannten Schulstufen und Schülern der HTBLA Kaindorf für die Erstellung der Aufnahmen, mit dem Ziel körperbehinderten Schülern mittels verbaler Steuerung sowohl die Nutzung herkömmlicher Lernsoftware, als auch die Diktation der geforderten Lehrplanziele der 1. und 2. Schulstufe zu ermöglichen. Falls dieses Projekt positiv abgeschlossen wird, könnten in Folgeprojekten ebenfalls Diktationsmodelle für die 3. und 4. Schulstufe entwickelt werden, die wesentlich komplexer sind, aber körperbehinderten Schülern in der Zukunft eine aktive schriftliche Kommunikation auf Basis des wichtigsten Grundwortschatzes ermöglichen könnten. Die Erstellung dieser Sprachmodelle ist nur mit aktiver Forschungsteilnahme von Schülern der 1. bis zur 4. Schulstufe der Volksschulen möglich, da es erforderlich ist den entsprechenden Wortschatz und die entsprechenden Satzstrukturen mit Kinderstimmen aufzunehmen und daraus die entsprechenden Sprachmodelle zu generieren. Das Endergebnis dieses Projektes soll den Schülern wieder durch eine verbal steuerbare Station an ihrer Schule zu Testzwecken und zur Überprüfung des Endergebnisses zur Verfügung gestellt werden. Weiters soll im Rahmen dieses Projektes damit begonnen werden, Simon Szenarien für die Steuerung von schulischen Lernsoftwareprodukten zu entwickeln, damit körperbehinderte Schüler die Möglichkeit haben dieselben Lernsoftwareprodukte wie Normalschüler für Ihren Lernprozess zu verwenden.

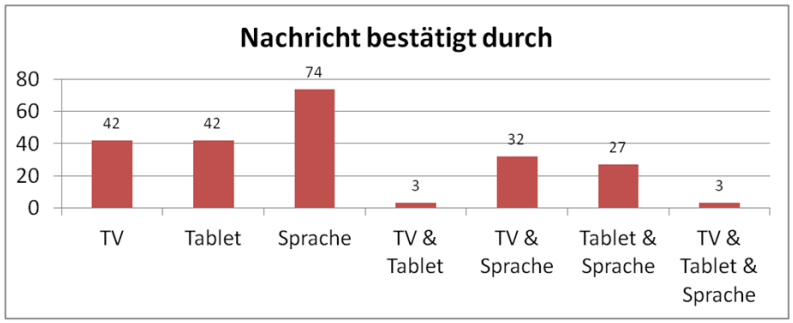


3.4 Perspektiven von Sprachsteuerung im Seniorenbereich

Neben den NSA finanziert arabischen Sprachmodellen, gibt es vor allem den Seniorenbereich, in dem sprachgesteuerte Lösungen in Zukunft gefordert sein werden. Die bisherigen Erfahrungen in Projekten und Erhebungen haben folgende Erkenntnisse ergeben. Eine Erhebung im Rahmen unseres ersten Benefit Projektes (Nationales Programm im AAL Bereich) ergab folgendes Ergebnis:



Motorische Störungen der Fein- und Grobmotorik sind bei Senioren dominant und treten frühzeitig ein. Die Hör- und Sehfunktionen sind durchschnittlich beeinträchtigt. Die Sprechfähigkeit bleibt am längsten erhalten und wird in der Regel spontan für die Steuerung eingesetzt. Die Auswertung der Userinteraktionen von Senioren im Rahmen des Forschungsprojektes „ibi“ ergab folgendes recht interessantes Ergebnis:



Anzahl der Nachrichteninteraktion über den Testungszeitraum. Hier ist es auffällig, dass die Nachrichteninteraktion gleichverteilt am Tablet als auch am TV Gerät eingeleitet wurde. Nach der Einleitung wurde vermehrt mit Hilfe der Sprachsteuerung navigiert und betätigt. Dies spiegelt sich auch an der hohen Zahl der Interaktionen (74), an denen die Sprachsteuerung involviert war wider.

3.5 Projektidee: Medizinisch diagnostische Sprachdatenbank

Vor allem bei der Entwicklung der Software SSC – Simon Sample Collector haben wir erstmals eine kleine diagnostische Sprachdatenbank von körperbehinderten Personen bestehend aus Basiswortschatz Simon Basisautonomie, Spontansprache und einem logopädischem Test erstellt. Diese Daten würden sie systematisch und wiederholt erhoben könnten einen immensen Forschungswert in der Diagnose und Therapie von körperbehinderten Klienten produzieren, und wäre sinnvoll als Forschungsprojekt wieder aufzunehmen.

Literatur

[1] Sachbericht zum Projekt Benefit Sprachsteuerung von IKT Anwendungen für Senioren (Download unter http://www.simon-listens.org/index.php?id=407)

[2] Sachbericht zum Projekt ASTROMOBILE (Download unter http://www.simon-listens.org/index.php?id=406)

[3] A. Zehetner, Keyword Spotting for Emergency, Master Thesis, Graz University of Technology, 2013

[4] A. Zehetner, M. Hagmüller, F. Pernkopf App112 Wake-Up-Word Spotting for Mobile Systems, Technical Report, Graz University of Technology, 2013