

Sprachtechnologie für Menschen mit Behinderung

Sarah Ebling
sarah.ebling@hfh.ch

Inhalt

- 1 Einleitung
- 2 Sprachtechnologie in assistierenden Technologien und als Teil von e-Accessibility
 - Spracherkennung
 - Sprachsynthese
- 3 Sprachtechnologie in sonderpädagogischen Interventionen
 - Spracherkennung
 - Sprachverstehen
 - Embodied Conversational Agents
- 4 Schluss

Gus, 13 Jahre alt

Gus: „You're a really nice computer.“

Siri: „It's nice to be appreciated.“

Gus: „You are always asking if you can help me. Is there anything you want?“

Siri: „Thank you, but I have very few wants.“

Gus: „O.K.! Well, good night!“

Siri: „Ah, it's 5:06 p.m.“

Gus: „Oh sorry, I mean, goodbye.“

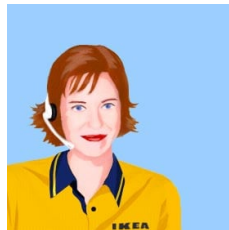
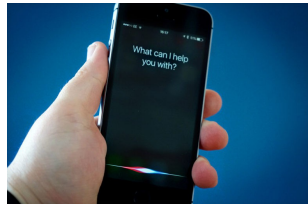
Siri: „See you later!“



Louie Chin

<https://www.nytimes.com/2014/10/19/fashion/how-apples-siri-became-one-autistic-boys-bff.html>

- ▶ Automatische Verarbeitung natürlicher Sprache
- ▶ Schnittstelle von Sprachwissenschaft und Informatik
- ▶ Beispiele:
 - ▶ Google Search
 - ▶ Google Translate
 - ▶ Rechtschreib- und Grammatikprüfung in Word
 - ▶ Diktiersoftware (z.B. Dragon Naturally Speaking)
 - ▶ Chatbots und Dialogsysteme: ELIZA, Anna (IKEA), **Siri**, Alexa, Cortana, ...



Künstliche Intelligenz

- ▶ Seit den 1950er-Jahren
- ▶ „a set of technologies that attempt to give computers some of the cognitive abilities of humans“ (Beals et al., 2015, S. 39)
- ▶ Teilbereiche:
 - ▶ **Sprachverstehen**
 - ▶ **Dialogmanagement**
 - ▶ Wissensrepräsentation
 - ▶ Emotionserkennung
 - ▶ **Maschinelles Lernen**
 - ▶ Objekterkennung
 - ▶ Robotik
 - ▶ ...

Künstliche Intelligenz: Und wo bleibt der Mensch?

„no computer system is capable of providing the same level of sophisticated and subtle translation that a qualified professional interpreter can“ (Huenerfauth and Hanson, 2009, S. 12)

„when it comes to language remediation, it turns out that computer technologies and human therapists are perfect complements“ (Beals et al., 2015, S. 65)

Sprachtechnologie für Menschen mit Behinderung

1. Sprachtechnologie in assistierenden Technologien und als Teil von e-Accessibility
2. Sprachtechnologie in sonderpädagogischen Interventionen

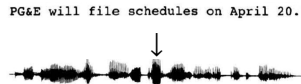
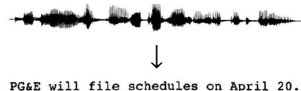
Ziele dieses Referats

1. Stand der Forschung/Entwicklung im Bereich Sprachtechnologie für Menschen mit Behinderung:
 - Was ist jetzt möglich?
 - Was wird in naher Zukunft möglich sein?
 - Was wird nie möglich sein?
2. Überblick über Methoden/Ansätze mit ihren Herausforderungen



Sprachtechnologien für Menschen mit Behinderung

- ▶ **Spracherkennung:** Umwandlung von gesprochener Sprache in Text
Spezialform: Gebärdenspracherkennung
 - ▶ Sprecherabhängig
 - ▶ Sprecherunabhängig
 - ▶ Sprecheradaptiv
- ▶ **Sprachsynthese:** Umwandlung von Text in gesprochene Sprache (*text to speech*, TTS)
Spezialform: Gebärdensprachsynthese
- ▶ **Sprachverstehen:** Erkennung der Struktur und Bedeutung eines Textes



Inhalt

- 1 Einleitung
- 2 Sprachtechnologie in assistierenden Technologien und als Teil von e-Accessibility
 - Spracherkennung
 - Sprachsynthese
- 3 Sprachtechnologie in sonderpädagogischen Interventionen
 - Spracherkennung
 - Sprachverstehen
 - Embodied Conversational Agents
- 4 Schluss

Inhalt

- 1 Einleitung
- 2 Sprachtechnologie in assistierenden Technologien und als Teil von e-Accessibility
 - Spracherkennung
 - Sprachsynthese
- 3 Sprachtechnologie in sonderpädagogischen Interventionen
 - Spracherkennung
 - Sprachverstehen
 - Embodied Conversational Agents
- 4 Schluss

Live-Untertitelung mit Respeaking

- ▶ z.B. Sportübertragungen, Diskussionssendungen beim Schweizer Fernsehen
- ▶ Respeaking: „a technique in which a respeaker listens to the original sound of a **live** programme or event and respeaks it, including **punctuation marks and some specific features** for the deaf and hard of hearing audience, to a **speech recognition software**“ (Romero-Fresco, 2011)
- ▶ Warum ist Respeaking notwendig?
 - ▶ Satzzeichen sprechen
 - ▶ Ebenmässigeres Sprechbild → näher an Trainingsmaterial (sprecherabhängige Spracherkennung)



Vollautomatische Live-Untertitelung?

- ▶ EU-Projekt Savas (bis 2014, <http://www.fp7-savas.eu/>)
- ▶ Herausforderungen:
 - ▶ Klassifikation in Redebeiträge, Musik, Hintergrundgeräusche, ...
 - ▶ Segmentierung der Untertitel
 - ▶ Sprecheridentifikation (z.B. Moderatoren)
 - ▶ Markierung von Sprecherwechseln
 - ▶ Interpunktion
 - ▶ Eigennamen
 - ▶ Erkennung von *code switching*
 - ▶ Editierung
- „speed of automatic subtitles is likely to cause viewers to miss most of the images“ (Romero-Fresco and Pérez, 2014)
- ▶ ...

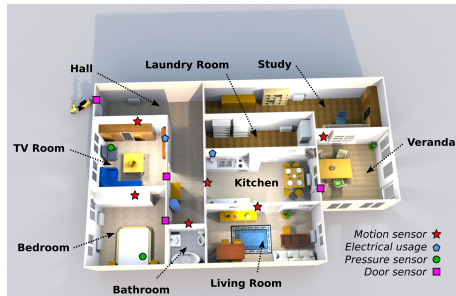


Automatische Erkennung dysarthrischen Sprechens

- ▶ Erkennungsgenauigkeit bei leichter bis mittelschwerer Dysarthrie: bis 90% mit kommerziellen Spracherkennungssystemen
- ▶ Schwere Dysarthrie: deutlich tiefer → sprecherabhängige Spracherkennung
- ▶ Herausforderungen (Hawley et al., 2013; Rudzicz, 2016):
 - ▶ Wenig Trainingsmaterial
 - ▶ z.T. sehr variables Trainingsmaterial
- ▶ Aufkommen neuer Datenbanken (UASpeech, TORGO)
- ▶ Laufende Initiative: **T**rainig **N**etwork on **A**utomatic Processing of **P**athological **S**peech (TAPAS, <https://www.tapas-etn-eu.org/>)

Spracherkennung für Active Assisted Living (AAL)

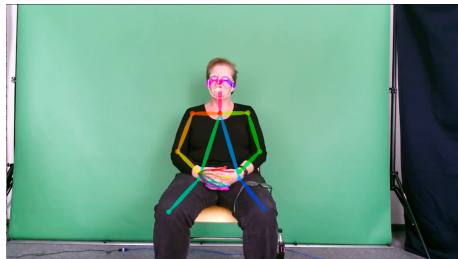
- ▶ Spracherkennung als Kommandoschnittstelle
- ▶ Befehle:
 - ▶ Notruf absetzen
 - ▶ Fensterläden auf/zurück
 - ▶ Lichtschalter ein/aus
 - ▶ ...
- ▶ Herausforderungen (Vacher et al., 2013, 2015):
 - ▶ Erkennung auf Distanz
 - ▶ Verschiedene Audiokanäle
 - ▶ Ständige Einsatzbereitschaft
 - ▶ Rücksicht auf Privatsphäre
 - ▶ Stimmen älterer Menschen können spezifische Merkmale aufweisen



(Palumbo et al., 2014)

Gebärdenspracherkennung

- ▶ Computerlinguistik und Computer Vision
- ▶ Zwei Schritte:
 1. Extraktion von Positionsmerkmalen für jeden Frame
 2. Zeitliche Modellierung
- ▶ Herausforderungen:
 - ▶ Allgemein: Lichtverhältnisse, Perspektive, Verdeckungen, Bewegungsunschärfe, ...
 - ▶ Gebärdensprachspezifisch: manuelle und nicht-manuelle Komponenten, wenig Trainingsmaterial, ...
- ▶ Bisher v.a. auf einzelnen Gebärden
- ▶ Erster Schritt: Tracking



Projekt SMILE

Inhalt

- 1 Einleitung
- 2 Sprachtechnologie in assistierenden Technologien und als Teil von e-Accessibility
 - Spracherkennung
 - Sprachsynthese
- 3 Sprachtechnologie in sonderpädagogischen Interventionen
 - Spracherkennung
 - Sprachverstehen
 - Embodied Conversational Agents
- 4 Schluss

Screen-Reader

- ▶ Seit Mitte der 1980er-Jahre
- ▶ Bekannte Screen-Reader: JAWS, VoiceOver (Accessibility Suite), NVDA
- ▶ Formantensynthese
- ▶ Merkmale (Zagler, 2014):
 - ▶ Synthesizer spricht zur blinden Person
 - ▶ Hohe Sprechgeschwindigkeit wichtig
 - ▶ Natürlichkeit der Stimme hat nicht oberste Priorität
 - ▶ Wiedergabe von Satzzeichen, Gross-/Kleinschreibung u.ä. teils von Bedeutung



(U5, IML Uni Bern)

Audiodeskription mit Sprachsynthese

Verwendung von Sprachsynthese: Studioaufnahme entfällt, Prooflistening orts- und zeitunabhängig

1. Sendematerial beschaffen
2. Sendematerial anschauen, Skript erstellen
3. Sprachsynthese
4. Prooflistening, Korrekturen
5. Zweiter Durchgang Sprachsynthese
6. Audiodeskription mit Ton der Sendung abmischen
7. Auf Sendedatei aufspielen



Inhalt

- 1 Einleitung
- 2 Sprachtechnologie in assistierenden Technologien und als Teil von e-Accessibility
 - Spracherkennung
 - Sprachsynthese
- 3 Sprachtechnologie in sonderpädagogischen Interventionen
 - Spracherkennung
 - Sprachverstehen
 - Embodied Conversational Agents
- 4 Schluss

Inhalt

- 1 Einleitung
- 2 Sprachtechnologie in assistierenden Technologien und als Teil von e-Accessibility
 - Spracherkennung
 - Sprachsynthese
- 3 Sprachtechnologie in sonderpädagogischen Interventionen
 - **Spracherkennung**
 - Sprachverstehen
 - Embodied Conversational Agents
- 4 Schluss

Spracherkennung in sonderpädagogischen Interventionen

- ▶ Transkription in Spontansprachanalysen
- ▶ Förderung/Therapie im Kontext von Dysarthrie (Ganzeboom et al., 2016; Krause et al., 2013; Martens et al., 2015)
Aktuelles Projekt: ISi-Speech (Frieg et al., 2017)
- ▶ Förderung/Therapie im Kontext von Aphasie (Beals et al., 2015)



(Ganzeboom et al., 2016)

Inhalt

- 1 Einleitung
- 2 Sprachtechnologie in assistierenden Technologien und als Teil von e-Accessibility
 - Spracherkennung
 - Sprachsynthese
- 3 Sprachtechnologie in sonderpädagogischen Interventionen
 - Spracherkennung
 - **Sprachverstehen**
 - Embodied Conversational Agents
- 4 Schluss

Sprachverstehen in sonderpädagogischen Interventionen

1. Morphologische Analyse
2. Wortartenanalyse
3. Syntaktische Analyse
4. Semantische Analyse

GrammarTrainer für Kinder/Jugendliche mit Autismus-Spektrum-Störung (Beals et al., 2015)

What does the boy want?

One word has the wrong ending. Change or delete the ending in red.

He wanted the girl to come over to his house and play with him

He wante the girl to come over to his house and play with him

Click or type the words to form the correct response. Then click 'Enter'.

a an and are at be boy boys came come comes coming did do does girl
they they'll they're to want wanted wanting wants will with you yourself

What will the boy want?

Change the noun phrase in red to a pronoun.

The boy will want the girl to help him

The boy will want the girl to help him

Click or type the words to form the correct response. Then click 'Enter'.

a an are at be boy boys did do does girl girls he help helped helping
wants will you yourself

What will the boy want?

The word in blue needs an article. Please type a, an, or the in the correct space.

He will want girl to come over to his house and play with him

He will want girl to come over to his house and play with him

Click or type the words to form the correct response. Then click 'Enter'.

a an and are at be boy boys came come comes coming did do does gi
they they'll they're to want wanted wanting wants will with you yourself

Textanalyse in sonderpädagogischen Interventionen

- ▶ Unterstützung Diagnose spezifische Sprachentwicklungsstörung (Solorio, 2013)
- ▶ Unterstützung Diagnose leichte kognitive Beeinträchtigung (Lehr et al., 2012)
- ▶ Unterstützung Diagnose Aphasie (Fraser et al., 2013, 2014)
- ▶ Unterstützung Bestimmung Sprachentwicklungsstand syntaktisch-morphologischer Bereich (Sagae et al., 2005)
- ▶ Auffinden von Texten geeigneten Niveaus
- ▶ Unterstützung Diagnostik PMT: Projekt „Computergestütztes Übungsprogramm für psychomotorische Diagnostik“ (https://www.hfh.ch/de/forschung/projekte/computergestuetztes_uebungsprogramm_fuer_psychomotorische_diagnostik/)

Inhalt

- 1 Einleitung
- 2 Sprachtechnologie in assistierenden Technologien und als Teil von e-Accessibility
 - Spracherkennung
 - Sprachsynthese
- 3 Sprachtechnologie in sonderpädagogischen Interventionen
 - Spracherkennung
 - Sprachverstehen
 - Embodied Conversational Agents
- 4 Schluss

Conversational Agents

- ▶ „computer programs engaging with human users in a conversation to assist, educate, or entertain“ (Suendermann-Oeft, 2014, S. 63)
- ▶ Alternative Bezeichnung: *Dialogsysteme*
- ▶ Beispiele für multimodale Dialogsysteme: Embodied Conversational Agents (ECAs)
 - ▶ „graphically embodied agents which aim to unite gesture, facial expression and speech to enable face-to-face communication with users“ (Amores et al., 2011, S. 334)
 - ▶ Computerlinguistik, Computergrafik
 - ▶ Herausforderungen (Lopez-Cozar et al., 2011):
 - ▶ Grössere kognitive Beanspruchung
 - ▶ Möglichkeit der Fehlerfortpflanzung
 - ▶ Höhere Rechnerleistung erforderlich

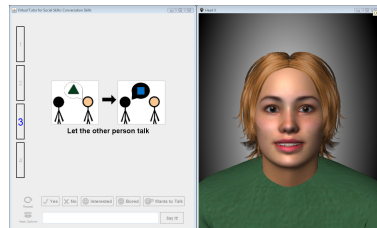


(Milne et al., 2011)

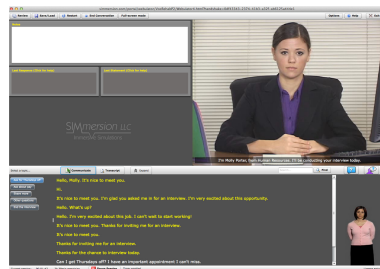
(Beals et al., 2015; Suendermann-Oeft, 2014)

ECAs in sonderpädagogischen Interventionen

- ▶ Autismus-Spektrum-Störung (Bosseler and Massaro, 2003; Milne et al., 2011; Smith et al., 2014; Tartaro and Cassell, 2008)
- ▶ Aphasie (Pompili et al., 2015; Teodoro et al., 2013)



(Milne et al., 2011)



(Smith et al., 2014)

Inhalt

- 1 Einleitung
- 2 Sprachtechnologie in assistierenden Technologien und als Teil von e-Accessibility
 - Spracherkennung
 - Sprachsynthese
- 3 Sprachtechnologie in sonderpädagogischen Interventionen
 - Spracherkennung
 - Sprachverstehen
 - Embodied Conversational Agents
- 4 Schluss

- ▶ Beitrag der Sprachtechnologie in sonderpädagogischen Interventionen: grosses Potenzial
- ▶ Evaluation sprachtechnologischer Systeme:
 - ▶ Sonderpädagogen
 - ▶ Benutzer



Vielen Dank



Q r 0 u (x 0 ~) ±

sarah.ebling@hfh.ch | ebling@cl.uzh.ch

Literatur I

- J. Gabriel Amores, Pilar Manchon, and Guillermo Perez. *Humanizing Conversational Agents: Indisys Practical Case Study in eHealth*, chapter 14, pages 312–334. Information Science Reference - Imprint of: IGI Publishing, Hershey, PA, 2011.
- Katharine Beals, Deborah Dahl, Ruth Fink, and Marcia Linebarger. *Speech and Language Technology for Language Disorders*. De Gruyter, Berlin, Boston, 2015.
- A. Bosseler and D. Massaro. Development and evaluation of a computer-animated tutor for vocabulary and language learning in children with autism. *Journal of Autism and Developmental Disorders*, 33(6):653–672, 2003.
- Sarah Ebling, Rosalee Wolfe, Jerry Schnepf, Souad Baowidan, John McDonald, Robyn Moncrief, Sandra Sidler-Miserez, and Katja Tissi. Synthesizing the finger alphabet of Swiss German Sign Language and evaluating the comprehensibility of the resulting animations. In *Proceedings of the 6th Workshop on Speech and Language Processing for Assistive Technologies (SLPAT)*, Dresden, Germany, 2015.
- Sarah Ebling, Sarah Johnson, Rosalee Wolfe, Robyn Moncrief, John McDonald, Souad Baowidan, Tobias Haug, Sandra Sidler-Miserez, and Katja Tissi. Evaluation of Animated Swiss German Sign Language Fingerspelling Sequences and Signs. In M. Antona and C. Stephanidis, editors, *Universal Access in Human-Computer Interaction*, LNCS, pages 1–13. Springer, 2017.
- Eleni Efthimiou, Stavroula-Evita Fotinea, Thomas Hanke, John Glauert, Richard Bowden, Annelies Braffort, Christophe Collet, Petros Maragos, and François Lefebvre-Albaret. Sign language technologies and resources of the Dicta-Sign project. In *Proceedings of the 5th LREC Workshop on the Representation and Processing of Sign Languages*, pages 37–45, Istanbul, Turkey, 2012.
- K. C. Fraser, J. A. Meltzer, N. L. Graham, C. Leonard, G. Hirst, S. E. Black, and E. Rochon. Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex*, 55:43–60, 2014.
- Kathleen Fraser, Frank Rudzicz, Naida Graham, and Elizabeth Rochon. Automatic speech recognition in the diagnosis of primary progressive aphasia. In *Proceedings of the Fourth Workshop on Speech Processing for Assistive Technologies*, pages 47–54, 2013.
- Hendrike Fried, Juliane Muehlhaus, Ute Ritterfeld, and Kerstin Bilda. Assistive technologien in der dysarthrietherapie. entwicklung des trainingssystems isi-speech als anwendungsbeispiel. 31:10–15, 05 2017.

Literatur II

- Mario Ganzeboom, Emre Yilmaz, Catia Cucchiaroni, and Helmer Strik. An asr-based interactive game for speech therapy. In *Proceedings of the 7th Workshop on Speech and Language Processing for Assistive Technologies (SLPAT 2016)*, pages 63–68. Association for Computational Linguistics, 2016.
- M.S. Hawley, S.P. Cunningham, P.D. Green, P. Enderby, R. Palmer, S. Sehgal, and P. O'Neill. A voice-input voice-output communication aid for people with severe speech impairment. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 21(1):23–31, 2013.
- Matt Huenerfauth and Vicki L. Hanson. Sign language in the interface: Access for Deaf signers. In Constantine Stephanidis, editor, *Handbook of Universal Access*. CRC Press, Boca Raton, FL, 2009.
- Michael Kipp, Quan Nguyen, Alexis Heloir, and Silke Matthes. Assessing the Deaf user perspective on sign language avatars. In *Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS)*, pages 107–114, Dundee, Scotland, 2011.
- Markus Krause, Jan Smeddinck, and Ronald Meyer. A digital game to support voice treatment for parkinson's disease. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 445–450, New York, NY, USA, 2013. ACM.
- Maiden Lehr, Emily Prud'hommeaux, Izhak Shafran, and Brian Roark. Fully automated neuropsychological assessment for detecting mild cognitive impairment. In *Proceedings of the 13th Annual Conference of the International Speech Communication Association*, pages 1039–1042, 2012.
- Ramon Lopez-Cozar, Zoraida Callejas, Gonzalo Espejo, and David Griol. *Enhancement of Conversational Agents by Means of Multimodal Interaction*, chapter 10, pages 223–252. Information Science Reference - Imprint of: IGI Publishing, Hershey, PA, 2011.
- Heidi Martens, Gwen Van Nuffelen, Tomas Dekens, Maria Hernández-Díaz Huici, Hector Arturo Kairuz Hernández-Díaz, Miet De Letter, and Marc De Bodt. The effect of intensive speech rate and intonation therapy on intelligibility in parkinson's disease. *Journal of Communication Disorders*, 58:91 – 105, 2015. ISSN 0021-9924. doi: <https://doi.org/10.1016/j.jcomdis.2015.10.004>. URL <http://www.sciencedirect.com/science/article/pii/S0021992415300113>.

Literatur III

- Marissa Milne, Martin Luerksen, Trent Lewis, Richard Leibbrandt, and David Powers. *Designing and Evaluating Interactive Agents as Social Skills Tutors for Children with Autism Spectrum Disorder*, chapter 2, pages 23–48. Information Science Reference - Imprint of: IGI Publishing, Hershey, PA, 2011.
- Filippo Palumbo, Jonas Ullberg, Ales Štívec, Francesco Furfari, Lars Karlsson, and Silvia Coradeschi. Sensor network infrastructure for a home care monitoring system. *Sensors*, 14(3), 2014.
- Anna Pompili, Cristiana Amorim, Alberto Abad, and Isabel Trancoso. Speech and language technologies for the automatic monitoring and training of cognitive functions. In *Proceedings of SLPAT 2015: 6th Workshop on Speech and Language Processing for Assistive Technologies*, pages 103–109. Association for Computational Linguistics, 2015. doi: 10.18653/v1/W15-5118. URL <http://www.aclweb.org/anthology/W15-5118>.
- P. Romero-Fresco and J. Martínez Pérez. Accuracy rate in live subtitling: The NER model. In J. Díaz Cintas and J. Neves, editors, *Media for All 4*. 2014.
- Pablo Romero-Fresco. *Subtitling through speech recognition: Respeaking*. St. Jerome Publishing, Manchester, 2011.
- Frank Rudzicz. *Clear Speech: Technologies that Enable the Expression and Reception of Language*. Morgan & Claypool, 2016.
- Kenji Sagae, Alon Lavie, and Brian MacWhinney. Automatic measurement of syntactic development in child language. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, ACL '05, pages 197–204, Stroudsburg, PA, USA, 2005. Association for Computational Linguistics.
- M. J. Smith, E. M. Ginger, K. Wright, M. A. Wright, J. L. Taylor, H. L. Boteler, D. Olsen, M. B. Bell, and M. F. Fleming. Virtual reality job interview training in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 44:2450–2463, 2014.
- Thamar Solorio. Survey on emerging research on the use of natural language processing in clinical language assessment of children. *Language and Linguistics Compass*, 7(12):633–646, 2013.
- David Suendermann-Oeft. Modern conversational agents. In J. Jähnert and C. Förster, editors, *Technologien für digitale Innovationen*. Springer Fachmedien, Wiesbaden, Germany, 2014.

Literatur IV

- A. Tartaro and J. Cassell. Playing with virtual peers: Bootstrapping contingent discourse in children with autism. In *Proceedings of the International Conference of the Learning Science*, Utrecht, The Netherlands, 2008. ACM Press.
- G. Teodoro, N. Martin, E. Keshner, J. Y. Shi, and A. I. Rudnicky. Virtual clinicians for the treatment of aphasia and speech disorders. *Virtual rehabilitation (IVVR)*, pages 158–159, 2013.
- Michel Vacher, Benjamin Lecouteux, Dan Istrate, Thierry Joubert, François Portet, and Mohamed Sehili adn Pedro Chahuara. Experimental evaluation of speech recognition technologies for voice-based home automation control in a smart home. In *Proceedings of the 4th Workshop on Speech and Language Processing for Assistive Technologies (SLPAT)*, pages 99–105, Grenoble, France, 2013.
- Michel Vacher, Sybille Caffiau, François Portet, Brigitte Meillon, Camille Roux, Elena Elias, Benjamin Lecouteux, and Pedro Chahuara. Evaluation of a context-aware voice interface for Ambient Assisted Living: Qualitative user study vs. quantitative system evaluation. *ACM Transactions on Accessible Computing*, 7(2), 2015.
- Wolfgang Zagler. Ambient Assisted Living, 2014. Lecture held at the Institute of Computational Linguistics of the University of Zurich, Switzerland (December 17, 2014).