

## Personalisierte Sprachdialogsysteme für Kinder im Fahrzeug

Yara Behrens<sup>1,2</sup>, Ulrich Heid<sup>1</sup>, Thomas Rohdenburg<sup>2</sup>

<sup>1</sup> *Universität Hildesheim, 31141 Hildesheim, E-Mail: {behrens,y,heid}@uni-hildesheim.de*

<sup>2</sup> *IAV GmbH, 38518 Gifhorn, E-Mail: thomas.rohdenburg@iav.de*

### Einleitung

Die rasante Verbreitung von Sprachassistenten in deutschen Haushalten führt dazu, dass auch Kinder immer häufiger mit diesen in Kontakt kommen. Dabei lassen sich häufig Kommunikationsprobleme beobachten, da die Hauptzielgruppe der Sprachassistenten weiterhin Erwachsene sind. Auch wenn sich inzwischen vereinzelt Inhalte an Kinder richten, wird doch oftmals außer Acht gelassen, dass diese von Erwachsenen abweichende Kommunikationsstrategien und Ansprüche an Inhalte haben.

Die Eingabemodalität Sprache macht Sprachdialogsysteme barrierefrei und ermöglicht so auch Kindern, noch vor Ausbildung von Fähigkeiten andere Nutzerschnittstellen zu bedienen, prinzipiell Zugang zu Informationen, Bildung und Unterhaltung. Eine Implementierung von Sprachdialogsystemen für Kinder im Fahrzeug ist ein interessanter Ansatz, diese in ihren Alltag zu integrieren.

Studien zeigen, dass die Personalisierung eines Sprachdialogsystems die Nutzerzufriedenheit bei Kindern verbessern kann. Dazu müssen sowohl technische als auch inhaltliche Aspekte berücksichtigt werden.

Ziel des Projekts war, einen Empfehlungskatalog zur Entwicklung personalisierter Sprachdialogsysteme für Kinder im Fahrzeug zu erstellen. Dazu wurden aktuelle Studienergebnisse mit eigenen Daten, welche mittels einer Umfrage sowie Interviews mit Kindern erhoben wurden, verglichen und erweitert. Im Fokus der Arbeit standen Kinder im Alter von 5-11 Jahren.

### Spracherkennung

In der Wachstumsphase verändern sich Kinderstimmen kontinuierlich. Diese Veränderungen sind alters- und geschlechtsabhängig. Das führt dazu, dass eine robuste Spracherkennung für Kinder andere Kriterien erfüllen muss als für Erwachsene. Zudem befindet sich das linguistische Wissen von Kindern in der Ausbildung, was zu Aussprache Fehlern, Pausen und Fehlstarts führt. Diese Faktoren erschweren eine robuste Spracherkennung für Kinder zusätzlich [1,2,3,4].

Mit zunehmendem Alter gleichen sich Formantfrequenzen, Satzdauer und Aussprache an Erwachsene an. Ab einem Alter von 11 Jahren bilden sich Unterschiede in Grund- und Formantfrequenzen zwischen den Geschlechtern heraus [3]. Die Frequenzen von Mädchen liegen dabei im Schnitt 10% höher als bei Jungen [1,2]. Fehlendes linguistisches Wissen führt zu Grammatik- und Aussprache Fehlern sowie Fehlstarts. Diese sind bei Kindern zwischen 8-10 Jahren zweimal häufiger als bei Kindern im Alter von 11-14 Jahren [5].

Studien zeigen außerdem, dass ein auf den Sprecher angepasstes akustisches Modell die Fehlerrate in der Spracherkennung um 45% reduzieren kann und die Erkennung von Sprechern ab 9 Jahren vergleichbar mit Erwachsenen ist [5].

### Sprachverstehen

Die Entwicklung des Sprachverstehens kann bei einer normalen Entwicklung des Kindes auf grammatikalischer Ebene bestimmten Altersgruppen zugeordnet werden. Dabei ist es essenziell, zwischen aktivem und passivem Wortschatz zu unterscheiden. Zu Beginn der Schulzeit beinhaltet das aktive Vokabular 3.000-5.000 Wörter, das passive 9.000-14.000 [6].

In Laufe der Grundschulzeit werden neue grammatikalische Komponenten der Kategorien *Wortschatz*, *Morphosyntax*, *Satzsemantik* und *Pragmatik* erlernt, z.B.:

*Wortschatz:* Abstrakte Begriffe, Idiome

*Morphosyntax:* Konjunktiv, Infinitivsätze, Passivsätze

*Satzsemantik:* Reflexivpronomen, komplexe Quantoren

*Pragmatik:* Argumentationsausdrücke, Konjunktionen, Ironie

Dabei zeigt sich ein Entwicklungssprung zwischen 6-8 Jahren, analog zur Entwicklung der eigenen Wahrnehmung sowie der Wahrnehmung der Umwelt [6,7]. Ein Sprachdialogsystem sollte diese Perspektivverschiebung durch angepasste Inhalte und Dialoge unterstützen.

### Kindgerechte Inhalte und Dialoge

Neuere Ergebnisse zeigen, dass Sprachdialogsysteme für Kinder nicht nur technisch, sondern auch inhaltlich an das jeweilige Alter angepasst werden müssen.

Dafür sollten Kinder in Altersgruppen eingeteilt werden. Kinder zwischen 4-6 Jahren testen ihre Grenzen aus, wollen alternative Wege ausprobieren und entwickeln ein Interesse an ihrer Umwelt. Die 6 bis 8-jährigen haben oft ein Gefühl von Kontrollverlust. Mit Beginn der Schulzeit verändern sich bestehende soziale Strukturen und viele Kinder sind erstmals längere Zeit außer Haus. Um Kontrolle über die neue Situation zurückzuerlangen, wollen sie bei neuen Aufgaben angeleitet werden, um Fehler zu vermeiden. Die 8 bis 10-jährigen testen erneut ihre Grenzen und eigene Wege aus. Zwischen 10-12 Jahren beginnen Kinder Unterschiede zwischen sich und anderen wahrzunehmen, was zu Identitätsverlust führen kann [7].

Weitere Studien zeigen, dass Kinder Schwierigkeiten haben, Anfragen an ein Sprachdialogsystem neu zu formulieren, wenn dieses Kontextinformationen benötigt. Durch fehlendes Wissen über Kommunikationsstrategien behalten sie bei

erneuten Anfragen die originale Satzstruktur bei, betonen Wörter anders, machen mehr Pausen im Satz oder sprechen lauter [8,9,10]. Ein Feedback vom System, an welcher Stelle die Anfrage nicht verstanden wurde bzw. das Herstellen von Kontexten aus früheren Anfragen, kann Kindern die Interaktion erleichtern [10]. Zudem sollten Antworten vereinfacht und an das Alter und den Wissensstand des Kindes angepasst werden [11]. Auch das Herstellen einer Synergie im Dialog zwischen Sprachdialogsystem und Kind kann die Interaktion positiv beeinflussen. Diese kann durch widerspiegeln des Kommunikationsstils des Kindes erreicht werden [9].

Mittlerweile gibt es erste, speziell für Kinder entwickelte Systeme [12,13,14]. Dazu gehören sowohl Sprachdialogsysteme für Kinder im Auto [12,13], Spracherkenner mit angepassten akustischen Modellen [14,15,16,17] sowie Amazon Skills [18] bzw. der *Echo Dot Kids* [19].

### Datenerhebung

Mithilfe einer Umfrage unter Erwachsenen und Interviews mit Kindern wurde ermittelt, wie Kinder Sprachdialogsysteme nutzen, welche Probleme in der Interaktion auftreten, ob ein personalisiertes Sprachdialogsystem im Auto gewünscht wird und welche Funktionen dieses haben sollte.

Die Umfrage diente als Grundlage für weiterführende Interviews mit Kindern und untersuchte die o.g. Fragen aus Sicht Erwachsener (mit und ohne Kinder). Es nahmen insgesamt 90 Personen teil. Dabei war der Anteil an Teilnehmern mit und ohne Kinder ungefähr gleich verteilt. Von den eingetragenen Kindern gehören 43 der Zielgruppe an.

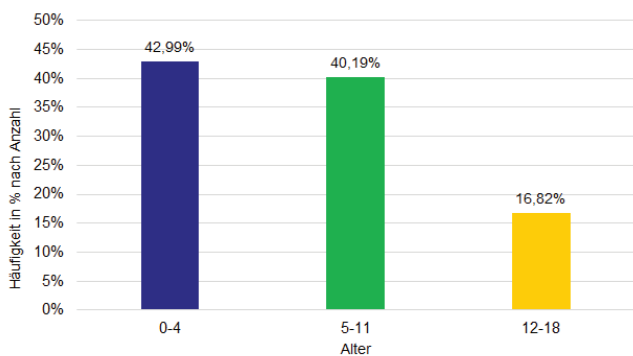


Abbildung 1: Prozentualer Anteil Kinder pro Altersgruppe.

Im Anschluss an die Umfrage wurden persönliche Interviews mit 14 Kindern durchgeführt. Davon waren 10 Jungen und 4 Mädchen. Sie wurden dabei zu ihren bisherigen Erfahrungen mit Sprachdialogsystemen, ihrem Nutzungsverhalten, Beschäftigungen bei Autofahrten sowie Lieblingsfiguren aus Hörspielen befragt. Letzteres dient der Untersuchung, welche Charaktereigenschaften und Stimmen Kinder favorisieren um daraus Rückschlüsse auf die Persönlichkeit eines Sprachdialogsystems ziehen zu können. Da Hörspiele einen rein auditiven Charakter aufweisen, eignen sie sich hier in besonderem Maße zur Evaluation, da Sprachdialogsysteme ebenfalls auditiv agieren.

Von den Interviews wurden Tonaufzeichnungen angefertigt, um das Kind während der Befragung nicht abzulenken und

um im Anschluss wichtige Aussagen notieren zu können. Den Eltern wurde freigestellt, ob sie während des Interviews anwesend sein wollten.

### Ergebnisse

Genutzt werden Sprachdialogsysteme von Kindern hauptsächlich um Musik oder Hörspiele zu hören oder zur Unterhaltung.

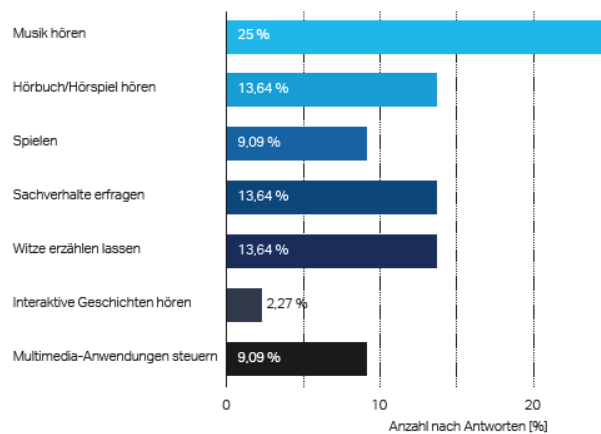


Abbildung 3: Nutzung von Sprachdialogsystemen durch Kinder (Teilnehmer mit Kindern, Mehrfachauswahl).

Dabei liegen die häufigsten Probleme in der Interaktion zwischen Kindern und Sprachdialogsystemen in der Spracherkennung bzw. fehlender Reaktion des Sprachdialogsystems. Auch falsche/unangemessene Antworten kommen häufig vor, welche die Kinder verunsichern.

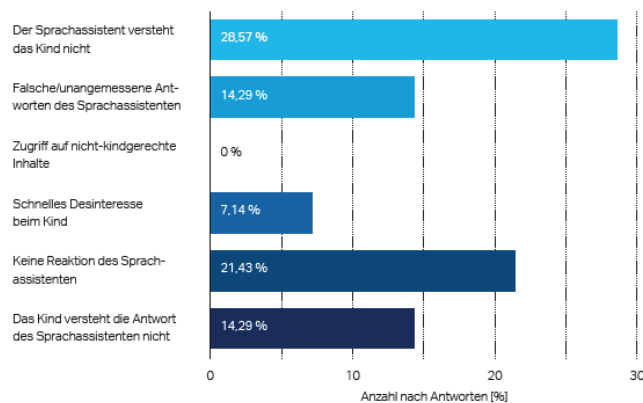
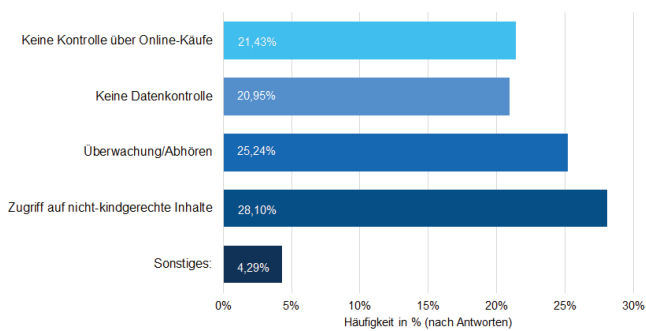


Abbildung 2: Herausforderungen für Kinder bei der Nutzung von Sprachdialogsystemen (Teilnehmer mit Kindern, Mehrfachauswahl).

Die Antworten zeigen, dass ein Zugriff auf nicht-kindgerechte Inhalte von Teilnehmern mit Kindern nicht beobachtet wurde. Interessant ist, dass in einer weiteren Frage bzgl. Bedenken in der Nutzung von Sprachdialogsystemen durch Kinder, dieser Punkt am häufigsten genannt wird. Eine Analyse der Ergebnisse zeigt, dass dieser Aspekt mehrheitlich von Teilnehmern ohne Kinder und/oder ohne Sprachdialogsystem im Haushalt genannt wird. Hier zeigt sich eine Diskrepanz zwischen erwarteten und tatsächlich beobachteten Problemen.



**Abbildung 3:** Bedenken bei der Nutzung von Sprachdialogsystemen durch Kinder (alle Teilnehmer, Mehrfachauswahl).

Insgesamt bewerten die Teilnehmer Sprachdialogsysteme als verbesserungswürdig. Dabei wünschen sie sich eine robustere Spracherkennung, altersangepasste Inhalte und Dialoge sowie Funktionen zur Kontrolle bzw. Sperrung von Inhalten.

Ein Sprachdialogsystem für Kinder im Auto kann sich die Mehrheit der Teilnehmer vorstellen. Dieses sollte das Kind während der Fahrt mit interaktiven Geschichten und Spielen beschäftigen und ihm bestenfalls dabei Wissen vermitteln. Als Hauptgrund gegen ein solches System im Auto wurden Fahrerablenkung und eine mögliche Reduktion der Kommunikation zwischen Passagieren angeführt. Daher müssen bei der Implementierung eines Sprachdialogsystems ins Fahrzeug geltende Richtlinien und Empfehlungen zur Fahrerablenkung berücksichtigt werden.

In den weiterführenden persönlichen Interviews mit Kindern wurden die Ergebnisse der Umfrage bestätigt und erweitert. Sprachdialogsysteme werden hauptsächlich für Musik/Hörspiele und zur Unterhaltung genutzt, wobei es häufig zu Fehlern in der Spracherkennung kommt, die die Kinder frustrieren und verunsichern. In den Antworten der Kinder zeigt sich aber auch, dass sie sich bewusst sind, dass es sich bei einem Sprachdialogsystem um ein computerbasiertes System handelt weshalb sie sich nach eigener Aussage toleranter gegenüber diesen Fehlern zeigen, als Erwachsene.

Eine Analyse der favorisierten Hörspiele und Figuren zeigt, dass Kinder Abenteuergeschichten präferieren, in denen sich eine Gruppe von Freunden für gemeinsame Ziele einsetzt. Zudem spiegeln genannte Figuren charakterliche Eigenschaften des Kindes. Daher sollten die Interessen und der Charakter eines Kindes eine zentrale Rolle bei der Personalisierung eines Sprachdialogsystems spielen. Kinder sollten zudem zwischen verschiedenen Stimmen und Figuren wählen können um optimal auf die Bedürfnisse der Kinder anpassbar zu sein. Auch sollte das System lernfähig sein, um angemessen auf neu erlernte Fähigkeiten und veränderte Interessen reagieren zu können und somit das Kind in seiner Entwicklung zu unterstützen.

Ein Großteil der Kinder äußerte den Wunsch nach einem Monitor, auf dem sie einen personalisierten Avatar des Sprachdialogsystems sehen und mit ihm interagieren können. Ein rein sprachliches System erscheint auf Basis der Antworten nicht interessant.

## Fazit

Die Ergebnisse der Studien sowie der eigenen Datenerhebung zeigen, dass ein Bedarf an personalisierten Sprachdialogsystemen für Kinder besteht. Dabei reicht es jedoch nicht aus, gleiche Inhalte und Funktionen für alle Altersgruppen anzubieten, wie es für Erwachsene häufig der Fall ist. Im Gegenteil müssen diese speziell auf das Alter, den Charakter und die Interessen des Kindes zugeschnitten und das System lernfähig sein. Auch die Spracherkennung muss an Kinder angepasst werden, da sie höhere Frequenzen aufweisen und es durch fehlendes linguistisches Wissen häufig zu Aussprachefehlern und längeren Pausen innerhalb eines Satzes kommt. Daraus resultiert eine Empfehlung, auf Kinder angepasste akustische Modelle und Grammatiken zu verwenden, um die Performance signifikant zu verbessern.

Auch sollten für Eltern Funktionen eingebaut werden, die es ihnen erlauben, bestimmte Inhalte für Kinder zu sperren und die Nutzung zu überwachen.

Bei einer Implementierung im Fahrzeug müssen geltende Richtlinien und Empfehlungen zu Fahrerablenkung beachtet werden, um diese zu minimieren und damit die Nutzerzufriedenheit zu optimieren.

Zudem sollte die Implementierung eines Monitors in Betracht gezogen werden, da die im Interview befragten Kinder den Wunsch äußerten, die Figur, mit der sie reden, auch sehen und mit ihr interagieren zu können.

## Literatur

- [1] Bennett, S.: Vowel formant frequency characteristics of preadolescent males and females. *The Journal of the Acoustical Society of America* 69(1) (1981), 231-238
- [2] Busby, P. A. & Plant, G. L.: Formant frequency values of vowels produced by preadolescent boys and girls. *The Journal of the Acoustical Society of America* 97(4) (1995), 2603-2606
- [3] Lee, S., Potamianos, A. & Narayanan, S.: Acoustics of children's speech: developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America* 105(3) (1999), 1455-1468
- [4] Huber, J. E. et al.: Formants of children, women, and men: the effects of vocal intensity variation. *The Journal of the Acoustical Society of America* 106(3) (1999), 1532-1542
- [5] Narayanan, S. & Potamianos, A.: Creating Conversational Interfaces for children. *IEEE Transactions on Speech and Audio Processing* 10(2) (2002), 65-78
- [6] Schulz, P.: Erstspracherwerb Deutsch: Sprachliche Fähigkeiten von Eins bis Zehn. Schneider-Verlag, Hohengehren, 2008
- [7] Gelman, D. L.: Design for Kids. Digital Products for Playing and Learning. Rosenfeld Media, Brooklyn, 2014

- [8] Lovato, S. B. & Piper, A. M.: "Siri, is this you?": Understanding Young Children's Interactions with Voice Input Systems. ACM Press, New York, 2015
- [9] Druga, S., Williams, R., Breazeal, C. & Resnick, M.: "Hey Google is it OK if I eat you?". ACM Press, New York, 2017
- [10] Yarosh, S. et al.: Children Asking Questions: Speech Interface Reformulations and Personification Preferences. Proceedings of the 17th ACM Conference on Interaction Design and Children (2018), 300-312
- [11] Lovato, S. B., Piper, A. M. & Wartella, E. A.: "Hey Google, Do Unicorns Exist?". ACM Press, New York, 2019
- [12] Broy, N. et al.: A Cooperative In-Car Game for Heterogeneous Players. ACM Press, New York, 2011
- [13] Gordon, M. & Breazeal, C.: Designing a virtual assistant for in-car child entertainment. ACM Press, New York, 2015
- [14] SoapBoxLabs Homepage, URL: <https://www.soapboxlabs.com>
- [15] KidSense.ai Homepage, URL: <https://kidsense.ai>
- [16] Moment Design Homepage, URL: <http://www.momentdesign.com/loree>
- [17] CogniToys Homepage, URL: <https://www.kickstarter.com/projects/cognitoys/cognitoys-internet-connected-smart-toys-that-learn>
- [18] BBC Kids Skill, URL: <https://www.amazon.co.uk/BBC-MEDIA-APPLICATIONS-TECHNOLOGIES-LIMITED/dp/B07BY9LKLG>
- [19] Honan, M: Amazon Created A Version Of Alexa Just For Kids, URL: <https://www.buzzfeednews.com/article/mathonan/amazon-alexa-for-kids#.ebZLxVrj8>