

„Ist jetzt ein guter Zeitpunkt?": Proaktive Ansprachen durch einen Sprachassistenten bei fahrfremden Tätigkeiten im Realverkehr

Lesley-Ann MATHIS¹, Daniela PIECHNIK², Carla Bernadette BUBECK¹,
Selina LAYER², Harald WIDLROITHER¹

¹ *Fraunhofer Institut für Arbeitswirtschaft und Organisation,
Nobelstraße 12, D-70569 Stuttgart*

² *Institut für Arbeitswissenschaft und Technologiemanagement, Universität Stuttgart,
Nobelstraße 12, D-70569 Stuttgart*

Kurzfassung: Automatisiertes Fahren ermöglicht Nutzenden das Ausführen fahrfremder Tätigkeiten, wie lesen oder relaxen während der Fahrt. Um das Wohlbefinden und Fahrerlebnis von Nutzenden zu steigern, bieten kontextorientierte Sprachassistenten zukünftig das Potenzial, proaktiv mit den Passagierinnen und Passagieren im Fahrzeug zu interagieren. Entscheidend für die Nutzerakzeptanz proaktiver Vorschläge ist die Wahl eines geeigneten Ansprachezeitpunkts. Die vorgestellte Realfahrstudie untersucht die Akzeptanz proaktiver Ansprachen bei unterschiedlichen Tätigkeiten der Nutzenden während der automatisierten Fahrt. Die Ergebnisse der Studie tragen zur Modellierung eines proaktiven Sprachassistenten im Fahrzeug bei, der mit Nutzenden zu geeigneten Zeitpunkten in den Dialog tritt.

Schlüsselwörter: Proaktivität, Sprachinteraktion, Automatisiertes Fahren, Fahrfremde Tätigkeit, Realfahrstudie, Wizard-of-Oz

1. Motivation und Stand der Forschung

Interaktion per Sprache gewinnt im Fahrzeug zunehmend an Bedeutung und wird von Nutzenden auch in automatisierten Fahrzeugen als bevorzugte Möglichkeit der Interaktion gesehen. Die Weiterentwicklung heutiger Sprachassistenten sieht vor, dass diese proaktive Vorschläge oder Angebote während der Fahrt machen können, um das Nutzungserlebnis der Fahrzeuginsassinnen und Fahrzeuginsassen zu steigern (Lugano 2017). So zeigen Studien für das manuelle Fahren, dass Nutzende proaktive Vorschläge zur Steigerung des Wohlbefindens wahrnehmen möchten, z. B. das Abspielen von Musik oder die Anleitung einer Achtsamkeitsübung (Koch et al. 2021). Eine Schlüsselfrage für die menschenzentrierte Gestaltung proaktiven Systemverhaltens ist, wann geeignete Zeitpunkte für Ansprachen der Nutzenden sind (Nothdurft et al. 2015). Voraussetzung hierfür ist die Kontexterfassung im Fahrzeug, die in zunehmend automatisierten Fahrzeugen sowohl durch Umfelderkennung als auch durch Innenraum-Monitoringsysteme ermöglicht wird.

Im Gegensatz zum manuellen Fahren zeigen erste Erkenntnisse zum automatisierten Fahren ab SAE-Level 3 (SAE International 2021), dass die Verkehrssituation für Nutzende eine untergeordnete Rolle für die Akzeptanz einer proaktiven Ansprache spielt (Mathis et al. 2023). Ein zu berücksichtigender Kontextfaktor hingegen ist die aktuelle Tätigkeit der Nutzenden, da automatisiertes Fahren ab SAE-Level 3 (SAE

International 2021) das Ausüben verschiedener fahrfremder Tätigkeiten (FFT) ermöglicht, wie z. B. Lesen oder Relaxen. In der vorgestellten Studie soll deshalb die Akzeptanz von proaktiven Ansprachen bei verschiedenen FFT untersucht werden.

2. Methode

Eine Realfahrstudie mit einem Wizard-of-Oz Versuchsfahrzeug wurde durchgeführt, bei der Teilnehmende proaktive Ansprachen erlebten. Das Versuchssetup wird im Folgenden näher beschrieben.

2.1 Studie im Realverkehr

Der Aufbau des Wizard-of-Oz Fahrzeugs mit insgesamt drei Fahrerplätzen (vorne rechts, vorne links und im Fond des Fahrzeugs) erlaubt es, das Fahrzeug von drei Plätzen zu steuern und somit verschiedene Automatisierungsstufen und Übergaben zu simulieren (Piechnik et al. 2023). Im Rahmen der Studie wurden auf einer Strecke von ca. 77 km (Fahrdauer: ca. 80 Min.) die SAE-Level 1–4 (SAE International 2021) simuliert. Für die proaktiven Ansprachen werden nur die Fahrtzeiten in SAE-Level 3 und 4 im Realverkehr analysiert, in denen die Versuchsleitung vorne rechts die Fahraufgabe innehatte, während die Versuchsperson auf dem Platz vorne links saß, auf dem üblicherweise die Fahrerin oder der Fahrer sitzt (siehe Abb. 1). Die Strecke beinhaltete sowohl Landstraßen- als auch Autobahnsequenzen und wurde vorab in mehrere Streckenabschnitte von jeweils ca. 8 bis 15 Minuten Dauer aufgeteilt.

Pro Streckenabschnitt wurden mögliche FFT auf einem Tablet in der Mittelkonsole des Fahrzeugs angezeigt. Diese beinhalteten die Tätigkeiten *Nachrichten am Handy* oder *ein Buch lesen*, *ein Telefongespräch führen*, *ein Handyspiel spielen*, *einen Kalendereintrag am Tablet vornehmen* oder *relaxen*. Des Weiteren war essen und trinken während der Fahrt jederzeit möglich.



Abbildung 1: Setup im Versuchsfahrzeug während der Fahrt im Realverkehr, mit der Versuchsperson auf dem Platz links vorne, hier beim Lesen eines Buches, und der Fahrerin rechts vorne.

Um möglichst natürliches Verhalten zu erzeugen, wurde den Teilnehmenden dabei freigestellt, ob, in welcher Reihenfolge und wie lange sie die jeweiligen Tätigkeiten ausführten. Neben der Ausführung der FFT wurden die Teilnehmenden pro definiertem Streckenabschnitt ein- bis zweimal von einem simulierten Sprachassistenten angesprochen. Die Teilnehmenden konnten insgesamt bis zu dreimal an der Studie teilnehmen, dabei blieb der Ablauf bei jeder Fahrt gleich.

2.2 Proaktive Ansprachen und Bewertung

Die proaktiven Ansprachen wurden mittels einer Softwareapplikation zur Dialogerstellung mit integrierter Text-to-Speech Funktion umgesetzt. Die Dialoge wurden durch eine weitere Person, die technische Versuchsleitung, im Fond des Fahrzeugs abgespielt, wobei jeder Dialog einem vordefinierten Schema folgte (siehe Abb. 2). Basierend auf der Studie von Semmens et al. (2019) startete jeder Dialog mit der Frage, ob gerade ein guter Zeitpunkt für eine Ansprache sei. Je nach Antwort der Versuchsperson wurde der entsprechende Dialogpfad ausgewählt. Wenn der Zeitpunkt als gut empfunden wurde, spielte die technische Versuchsleitung einen proaktiven Vorschlag ab, z. B. bat der Assistent seine Unterstützung bei der Parkplatzsuche am Zielort an. Jeder Dialog schloss mit der Frage, warum der jeweilige Zeitpunkt als gut bzw. schlecht für die Ansprache empfunden wurde. Alle Antworten der Versuchsperson wurden aufgezeichnet und die beobachtete FFT in der Softwareapplikation durch die Versuchsleitung dokumentiert. Um den Einfluss der Verkehrssituation und äußerer Ablenkung möglichst gering zu halten, löste die Versuchsleitung proaktive Ansprachen nur bei konstanter Fahrt auf gerader Strecke aus.

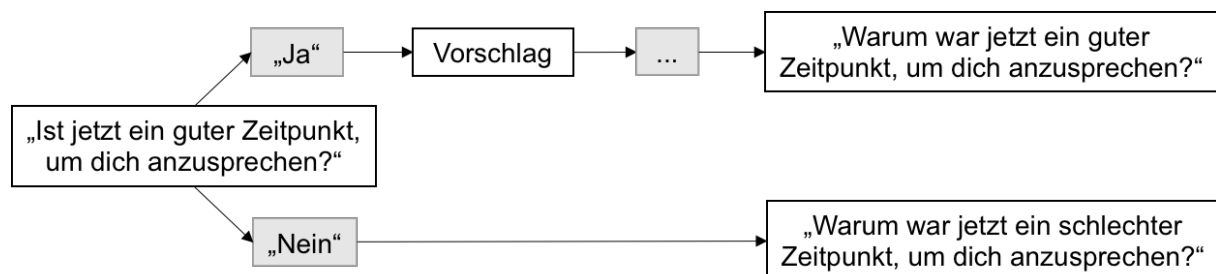


Abbildung 2: Schematischer Ablauf der proaktiven Ansprachen mit Systemausgaben (schwarz umrandet) und erwarteten Antworten der Versuchsperson (grau hinterlegt).

Als in-situ Bewertung der Ansprachen wird das Akzeptieren oder Ablehnen des proaktiven Vorschlags durch die Versuchsperson sowie die jeweilige Begründung ausgewertet. Darüber hinaus erhielten die Teilnehmenden nach der Fahrt einen Fragebogen, um die Akzeptanz von proaktiven Ansprachen bei verschiedenen FFT retrospektiv abzufragen. Dabei wurden die in der Studie untersuchten FFT mit der Frage „Wie geeignet empfinden Sie es, durch einen Sprachassistenten bei den folgenden Tätigkeiten angesprochen zu werden (vorausgesetzt, Sie dürfen diese Tätigkeit beim aktuellen Automatisierungslevel ausführen)?“ anhand einer 5-stufigen Likert-Skala von 1 (ungeeignet) bis 5 (geeignet) bewertet.

3. Ergebnisse

Insgesamt nahmen acht Personen an der Studie teil (5 männlich, 3 weiblich). Im Durchschnitt waren die Teilnehmenden $M = 36.6$ Jahre alt ($SD = 18.5$). Drei Teilnehmende nahmen an einem Termin, fünf Teilnehmende an drei Terminen der Studie teil.

3.1 In-situ Bewertung der Ansprachen bei FFT

Für die Analyse wurden $n = 92$ Ansprachen ausgewertet, nachdem dreizehn Ansprachen aufgrund unverständlicher Antworten oder technischer Probleme ausgeschlossen wurden. Die Teilnehmenden bewerteten 75 Prozent der Ansprachen insgesamt als gute Zeitpunkte, 25 Prozent als schlechte Zeitpunkte. Abb. 3 zeigt die relativen Häufigkeiten für positive Ansprachen pro beobachtete FFT, d. h. höhere Werte stellen eine häufigere positive Bewertung des Ansprachezeitpunkts dar. Ausschließlich gut wurden Ansprachen beim Hinausschauen bewertet, auch beim Essen sowie Relaxen zeigt sich eine hohe Zustimmung. Versuchsperson 7 begründete den geeigneten Zeitpunkt beim Relaxen folgendermaßen: „Weil ich mich entspannt habe und sonst nichts zu tun hatte.“

Ein heterogeneres Bild zeigt sich bei den Aktivitäten *Spielen*, *Lesen* und *Arbeiten* (*Kalendereintrag*). Beispielsweise empfand Versuchsperson 5 die Ansprache beim Spielen negativ: „Weil ich beim Spielen gestört werde. Ich kann verlieren“. Hingegen bewertete Versuchsperson 4 die Ansprache positiv: „Weil ich gerade ein Spiel gespielt habe am Handy, bei dem ich mich nicht konzentrieren muss.“ Bei der Tätigkeit *Lesen* zeigen die Begründungen der Teilnehmenden, dass entscheidend war, was gelesen wurde: Ansprachen beim Lesen am Handy ($n = 7$) oder Nachrichten lesen ($n = 3$) wurden wiederholt bei der Begründung für gute Zeitpunkte genannt. Hingegen wurden Ansprachen beim Lesen eines Buches ($n = 7$) häufig als schlecht wahrgenommen. Als uneingeschränkt gut wurden außerdem die Ansprachen beim Wechsel der Aktivität empfunden, so begründete Versuchsperson 2: „Ich hab’ gerade aufgehört mit Lesen und hatte Zeit.“

2.+ 3. Termin (5 P.)	100%	83%	50%	62%	50%	100%	0%		100%
1. Termin (8 P.)	100%	67%	60%	43%	50%			0%	100%
Gesamt	100%	78%	54%	54%	50%	100%	0%	0%	100%
	Hinausschauen (n=29)	Relaxen (n=9)	Spielen (n=13)	Lesen (n=35)	Kalendereintrag (n=4)	Essen (n=1)	Trinken (n=1)	im Gespräch (n=1)	Wechsel Aktivität (n=20)

Abbildung 3: Positive Bewertung in % der Ansprachen pro beobachtete FFT und aufgeteilt auf den 1. Termin sowie Wiederholungstermine (2.+3.Termin), inkl. Gesamtzahl der Ansprachen (n) pro beobachtete FFT

Die bisherigen Daten lassen keine Rückschlüsse darauf zu, ob sich ein wiederholtes Erleben der proaktiven Ansprachen bei der zweiten und dritten Fahrt positiv oder negativ auf die Bewertung der Ansprachezeitpunkte im Allgemeinen oder bei bestimmten FFT auswirkt.

3.2 Retrospektive Bewertung von Ansprachen bei FFT

The retrospektive Bewertung von Tätigkeiten auf einer Skala von 1 (ungeeignet) bis 5 (geeignet) bestätigt die positive Bewertung für die Tätigkeiten *Hinausschauen* ($M = 4.8$, $SD = 0.5$) und *Relaxen* ($M = 4.6$, $SD = 0.5$) sowie beim Wechsel der Aktivität ($M = 4.6$, $SD = 0.7$) für proaktive Ansprachen. Hingegen werden die Tätigkeiten *Arbeiten* ($M = 2.1$, $SD = 1.0$) oder *im Gespräch* ($M = 2.3$, $SD = 1.3$) als eher ungeeignet empfunden (siehe Abb. 4). Dazwischen liegen die Bewertungen für die Tätigkeiten *Essen* ($M = 3.3$, $SD = 1.5$), *Trinken* ($M = 3.0$, $SD = 1.2$) und *Spielen* ($M = 3.3$, $SD = 1.3$) mit neutral bis eher geeignet und für die Aktivität *Lesen* ($M = 2.6$, $SD = 0.7$) mit neutral bis eher ungeeignet.

Bei den offenen Kommentaren wiesen zwei Teilnehmende darauf hin, dass nicht nur die Tätigkeit an sich, sondern auch die Zeit bis zum Fahrtziel berücksichtigt werden sollte, da sie eine Ansprache kurz vor Ziel grundsätzlich nicht geeignet fänden.

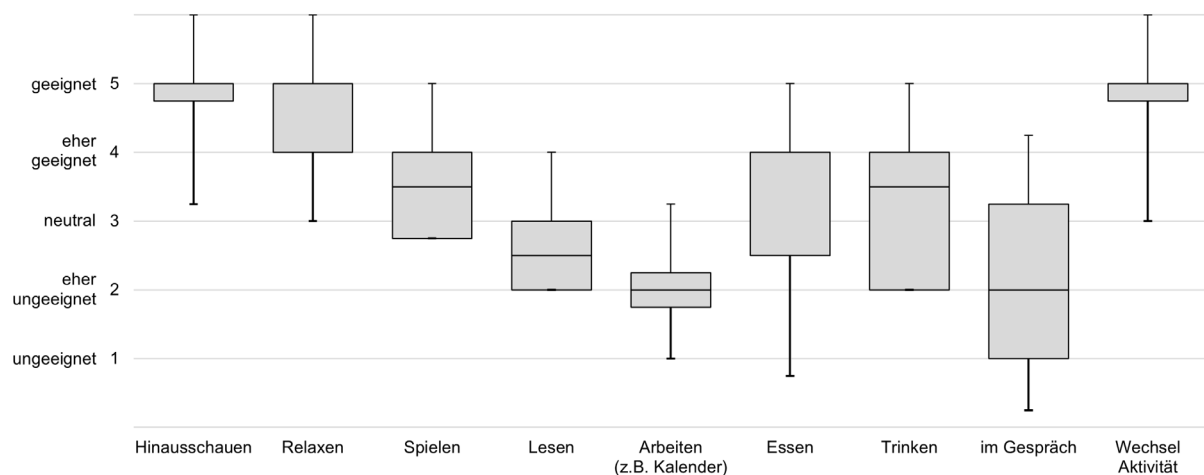


Abbildung 4: Boxplots der subjektiven Bewertung der 8 Teilnehmenden, wie (un)geeignet proaktive Ansprachen bei verschiedenen Tätigkeiten auf einer Likert Skala von 1 (ungeeignet) bis 5 (geeignet) empfunden werden.

4. Diskussion und Ausblick

Mittels einer Datenerhebung im Realverkehr wurde in der beschriebenen Studie untersucht, bei welchen FFT proaktive Ansprachen eines Sprachassistenten als geeignet oder ungeeignet empfunden werden. Die Ergebnisse der in-situ und retrospektiven Bewertung zeigen auf, dass sich Situationen, in denen die Passagierinnen und Passagiere keiner aktiven Aufgabe nachgehen, wie beim Hinausschauen, Relaxen oder beim Wechsel der Tätigkeit, am besten für proaktive Ansprachen eignen. Hingegen sind proaktive Ansprachen ungeeigneter, wenn sich Nutzende im Gespräch befinden oder einer Tätigkeit wie dem Arbeiten oder Lesen eines Buches nachgehen, die Konzentration erfordert. Das bestätigt die Ergebnisse einer Studie von Cha et al. (2020), die proaktive Ansprachen eines Sprachassistenten bei Tätigkeiten im häuslichen Umfeld untersuchte.

Des Weiteren ist für Ansprachen zu geeigneten Zeitpunkten im Fahrzeug nicht nur die Tätigkeit an sich relevant: Bei der Lesetätigkeit sollten Insassen-Monitoringssysteme differenzieren können, ob Nachrichten am Handy oder ein Buch gelesen

werden. Aktuell wird daran geforscht, mittels Künstlicher Intelligenz und der Fusion verschiedener Sensorquellen eine adaptive Interaktion mit Nutzenden zu ermöglichen (Diederichs et al. 2022). Durch die Kontexterkenkung können proaktive Sprachassistenten in Zukunft beim automatisierten Fahren z. B. rechtzeitig auf eine kurvige Strecke hinweisen, sodass Nutzende ihre Lesetätigkeit unterbrechen und Reisekrankheit vermeiden können.

Limitationen der Studie sind vor allem in der kleinen Stichprobe und einer unausgewogenen Anzahl von Ansprachen bei verschiedenen FFT zu sehen. Da die Studie ein möglichst natürliches Verhalten der Teilnehmenden anstrebt, bestehen große Unterschiede in der Häufigkeit der ausgeführten Tätigkeiten, sodass die Ergebnisse in zukünftigen Erhebungen validiert werden sollen. Auch soll untersucht werden, ob eine wiederholte Interaktion mit dem proaktiven Sprachassistenten die Bewertung der Nutzenden beeinflusst.

5. Literatur

- Cha N, Kim A, Park CY, Kang S, Park M, Lee J-G, Lee S, Lee U (2020) Hello There! Is Now a Good Time to Talk? Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 4: 1–28.
- Diederichs F, Wannemacher C, Faller F, Mikolajewski M, Martin M, Voit M, Widlroither H, Schmidt E, Engelhardt D, Rittger L, Hashemi V, Sahakyan M, Romanelli M, Kiefer B, Fäßler V, Rößler T, Großerüschkamp M, Kurbos A, Bottesch M, Immoor P, Engeln A, Fleischmann M, Schweiker M, Pagenkopf A, Mathis L-A, Piechnik D (2022) Artificial Intelligence for Adaptive, Responsive, and Level-Compliant Interaction in the Vehicle of the Future (KARLI). In: Stephanidis C, Antona M, Ntoa S (Hrsg.) HCI International 2022 Posters 164–171. Springer International Publishing, Cham.
- Koch K, Mishra V, Liu S, Berger T, Fleisch E, Kotz D, Wortmann F (2021) When Do Drivers Interact with In-Vehicle Well-being Interventions? An Exploratory Analysis of a Longitudinal Study on Public Roads. Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies 5: 1–30.
- Lugano G (2017) Virtual assistants and self-driving cars. In: Proceedings of the 15th International Conference on ITS Telecommunications (ITST) 1–5.
- Mathis L-A, Werner K, Schmidt GJ (2023) Proaktive Sprachassistenten in automatisierten Fahrzeugen. In: Hölzle K, Kreimeyer M, Roth D, Maier T, Riedel O (Hrsg.) Stuttgarter Symposium für Produktentwicklung SSP 2023, 488–499. Fraunhofer IAO, Stuttgart.
- Nothdurft F, Ultes S, Minker W (2015) Finding Appropriate Interaction Strategies for Proactive Dialogue Systems – An Open Quest. In: K. Jokinen, M. Vels (Hrsg.) Proceedings of the 2nd European and the 5th Nordic Symposium on Multimodal Communication 73–80. Linköping University Electronic Press.
- Piechnik D, Mathis L-A, Diederichs F, Lerch D, Martin M, Widlroither H (2023) Technical setup of a Wizard-of-Oz vehicle for on-road AI data collection. In: 15th ITS European Congress Lisbon, Portugal, Lisbon, Portugal.
- SAE International, Society of Automotive Engineers (2021) J3016 Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. SAE International – Recommended Practice.
- Semmens R, Martelaro N, Kaveti P, Stent S, Ju W (2019) Is Now A Good Time? In: Brewster S, Fitzpatrick G, Cox A, Kostakos V (Hrsg.) CHI 2019. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, May 4-9, 2019, Glasgow, Scotland UK 1–12. ACM, New York, NY.

Danksagung: Diese Studie fand im Rahmen des Projekts KARLI („Künstliche Intelligenz für eine adaptive, responsive und levelkonforme Interaktion mit dem Fahrzeug der Zukunft“), gefördert durch das Bundesministerium für Wirtschaft und Klimaschutz, statt.



Gesellschaft für Arbeitswissenschaft e.V.

Arbeitswissenschaft in-the-loop

**Mensch-Technologie-Integration
und ihre Auswirkung auf Mensch,
Arbeit und Arbeitsgestaltung**

70. Kongress der
Gesellschaft für Arbeitswissenschaft e.V.

Institut für Arbeitswissenschaft und
Technologiemanagement IAT
Universität Stuttgart

In Zusammenarbeit mit dem Fraunhofer-Institut für
Arbeitswirtschaft und Organisation IAO

06. – 08. März 2024

GfA-Press

Bericht zum 70. Arbeitswissenschaftlichen Kongress vom 06. – 08. März 2024

Institut für Arbeitswissenschaft und Technologiemanagement (IAT), Universität Stuttgart

In Zusammenarbeit mit: Fraunhofer-Institut für Arbeitswirtschaft und Organisation (IAO), Stuttgart

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.

Sankt Augustin: GfA-Press, 2024

ISBN 978-3-936804-34-8

NE: Gesellschaft für Arbeitswissenschaft: Jahresdokumentation

Als Manuskript zusammengestellt. Diese Jahresdokumentation ist nur in der Geschäftsstelle (s. u.) erhältlich.

Alle Rechte vorbehalten.

© **GfA-Press, Sankt Augustin, Schriftleitung: Prof. Dr. Rolf Ellegast**

im Auftrag der Gesellschaft für Arbeitswissenschaft e.V.

Ohne ausdrückliche Genehmigung der Gesellschaft für Arbeitswissenschaft e.V. ist es nicht gestattet:

- den Kongressband oder Teile daraus in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) zu vervielfältigen,
- den Kongressband oder Teile daraus in Print- und/oder Nonprint-Medien (Webseiten, Blog, Social Media) zu verbreiten.

Die Verantwortung für die Inhalte der Beiträge tragen alleine die jeweiligen Verfasser; die GfA haftet nicht für die weitere Verwendung der darin enthaltenen Angaben.

Geschäftsstelle der GfA

Simone John, Tel.: +49 (0)30 1300-13003, Alte Heerstraße 111, D-53757 Sankt Augustin

info@gesellschaft-fuer-arbeitswissenschaft.de · www.gesellschaft-fuer-arbeitswissenschaft.de

Screen design und Umsetzung

© 2024 fröse multimedia, Frank Fröse,

office@internetkundenservice.de, www.internetkundenservice.de