

## Editorial

Noch bevor viele von uns in eine hoffentlich erholsame Sommerpause gehen, erhalten Sie nun den mittlerweile dritten Newsletter der ITG-Fachausschüsse Sprachakustik (4.3) und Sprachverarbeitung (4.4). Eine ganze Reihe von Projekten und Journalartikeln sind es diesmal, vielleicht gibt es künftig auch noch mehr Links zu den (sicherlich vorhandenen) Dissertationen? Gerade die drohen nämlich oftmals unterzugehen und bieten doch eine hervorragende Lektüre und Darstellung innovativer Fachthemen... Bitte geben Sie unseren Newsletter weiter an interessierte Leser oder Personen aus Industrie und Wissenschaft, die ggfs. auch gerne selbst kleine Beiträge schreiben möchten. Wir wünschen Ihnen eine schöne Sommerzeit!

Ihr Tim Fingscheidt & Reinhold Häb-Umbach

Sie wünschen ein Abo oder haben einen Beitrag? Sehr gerne! Bitte melden Sie sich einfach [per Email](#) unter Hinweis darauf, ob Sie nur Abonnent, oder auch möglicher Autor sein möchten!

## Projekte und Aktivitäten



• Das dreijährige EU-Projekt [Embodied Audition for Robots \(EARS\)](#) (koordiniert von Walter Kellermann und Heiner Löllmann von der Friedrich-Alexander Universität Erlangen-Nürnberg, mit den Partnern Imperial College, London; Ben-Gurion Universität Beersheva; Humboldt Universität, Berlin; INRIA Grenoble; SoftBank Robotics, Paris) endete zum 31. Dezember 2016. Die entwickelten neuartigen Algorithmen für einen natürlichen Sprachdialog zwischen Mensch und humanoiden Robotern in realen akustischen Umgebungen beziehen „robot vision“ und „natürliche“ Bewegungen des Roboters mit ein. Nach mehreren [öffentlichen Demonstrationen](#) sind die Ergebnisse auch anhand von [Konferenz- und Zeitschriftenartikeln](#) und [Videos](#) nachzuvollziehen.

• Im März fand die diesjährige [DAGA-Konferenz](#) in Kiel statt. Etwa 1150 Teilnehmer hatten an vier Tagen die Wahl zwischen ca. 500 Vorträgen und konnten sich darüber hinaus ein Bild der neuesten Produkte aus dem Bereich der Akustik (vorgestellt von ca. 50 Ausstellern) machen. Ein erstmalig angebotenes (und sehr gut angenommenes) begleitendes Sportprogramm und die „[Nacht der Akustik](#)“ (mit zwölf allgemeinverständlichen Vorträgen für die Öffentlichkeit) rundeten das Konferenzprogramm ab.

• Ein Spracherkennungssystem für intelligente Wohnungen, mit dem auch Internetanwendungen robust und natürlich-sprachlich bedient werden können: Das entwickeln die Partner aus Griechenland, Italien und Deutschland im EU-Projekt „[LISTEN](#).“ Wichtige Bestandteile sind ein streichholzschachtelgroßes Mikrofon-Array sowie eine multilinguale Spracherkennung in Echtzeit. Beteiligte Partner sind der Heidelberger Sprachtechnologiespezialist EML European Media Laboratory GmbH, die RWTH Aachen, Cedat85 (Italien) und die griechische Foundation for Research and Technology Hellas (FORTH), die auch die Projektkoordination durchführt. Beim Projekttreffen auf Kreta testete die EU-Projektbeauftragte Alina Suhetzki „LISTEN“ persönlich [\[Link zum Video\]](#). „LISTEN“ ist eine Marie Skłodowska-Curie-Maßnahme für den Personalaustausch im Bereich Forschung und Innovation (RISE) im Rahmen von „Horizont 2020“ und läuft noch bis 2019.

• Die Professur *Kognitive Systeme* am Institut für Akustik und Sprachkommunikation startet im März 2017 ein neues BMBF-Förderprojekt (3 Jahre Laufzeit, zusammen mit der Fa. Linguwerk GmbH, Dresden) zur Entwicklung eines optoelektronischen Messsystems zur Steuerung interaktiver logopädischer Übungen in der Schlaganfalltherapie, sowie im Juni 2017 ein DFG-Projekt (3 Jahre Laufzeit) zur Untersuchung der intrinsischen Geschwindigkeiten von Artikulatoren beim Sprechen. Die Leitung der Projekte hat Jun.-Prof. Peter Birkholz.

• Im Fachgebiet *Quality & Usability* der TU Berlin ist das DFG-Projekt „[Simulation von Konversationsverhalten bei gestörter Telefonübertragung](#)“ gestartet. In diesem Projekt sollen Konversationen von der pragmatischen Ebene bis hin zum Sprecherwechsel simuliert werden. Dabei werden die Auswirkungen von Latenz und Paketverlust auf die Sprachverständlichkeit und den Sprecherwechsel modelliert. Mit Hilfe dieser Simulation soll eine Schätzung der Konversationsqualität vorgenommen werden.

- Gemeinsam mit der Firma [eye square GmbH](#) aus Berlin arbeitet das Institut für Nachrichtentechnik der Technischen Universität Braunschweig an der Entwicklung einer automatischen Beratung und Werbung im Online-Shopping. Ziel in diesem zweijährigen Förderprojekt des BMWi ist eine intelligente und auf den Kunden individuell angepasste Produktauswahl des Online-Sortiments, welche vom Wunsch und dem Geschmack des Kunden gesteuert wird. Die TU Braunschweig entwickelt in diesem Zusammenhang einen speziell auf diese Anwendung angepassten akustischen Emotionserkenner für mobile Endgeräte.

### Dissertationen

- Florian Heese, "[Speech Signal Enhancement by Information Combining](#)", RWTH Aachen University (P. Vary). The best of several speech and noise estimates are merged depending on their reliabilities. This method is called "information combining". The estimates are provided by a new robust statistical noise tracker and by a codebook driven speech/noise estimation algorithm allowing to trace highly non-stationary processes. The achieved quality opens up new perspectives to reduce the gap between the conflicting goals of high noise attenuation, low speech distortion, and no musical tone artifacts.

### Journalartikel

- F. Köster, D. Guse, S. Möller: [Identifying Speech Quality Dimensions in a Telephone Conversation](#)  
In this article, an approach towards analyzing speech quality in a conversational situation by dividing a conversation into three separate phases and identifying corresponding quality-relevant perceptual dimensions is presented. The identified dimensions can be combined for the overall quality assessment and may separately be used to diagnose the technical reasons of quality degradations.
- S. Elshamy, N. Madhu, W. Tirry, T. Fingscheidt: [Instantaneous A Priori SNR Estimation by Cepstral Excitation Manipulation](#)  
The article presents a new method for *a priori* SNR estimation by exploiting the source-filter model and enhancing the excitation signal, thereby allowing to achieve a higher noise suppression without further decreasing the speech component quality.
- P. Thüne, G. Enzner: [Maximum-Likelihood Approach with Bayesian Refinement for Multichannel-Wiener Postfiltering](#)  
This paper provides a perspective on multichannel Wiener filtering in time-varying conditions. We propose rigorous statistical estimation of the spec-

tral enhancement part of the filter, based merely on the available microphone signals and a minimum-informative prior of the ideal Wiener filter.

- F. Köster, F. Schiffner, S. Möller, L. Malfait: [Towards Degradation Decomposition for Voice Communication System Assessment](#)  
This article presents the current development of degradation decomposition tools for the assessment of voice communications. While MOS does not provide information on the cause of issues, work items are currently active within ITU-T to provide the industry with means to understand the cause of lower scores by perceptual or technical degradation decompositions. In this context, in sum four experiments using a potential diagnostic test method discussed within ITU-T are conducted, reviewed, and analyzed.

### Tagungen (nach Paper Deadline sortiert)

[Interspeech](#) 20.-24.08.2017 in Stockholm, Schweden, [keine Einreichungen mehr möglich]  
[EUSIPCO](#) 28.08.-02.09.2017 auf Kos, Griechenland, [keine Einreichungen mehr möglich]  
[MLSP](#) 25.-28.09.2017 in Tokyo, Japan [keine Einreichungen mehr möglich]  
[ASRU](#) 16.-20.12.2017 in Okinawa, Japan, **Paper Deadline 29.06.2017** [CfP]  
[ICASSP](#) 22.-27.04.2018 in Seoul, Korea, **Paper Deadline 27.10.2017** [CfP]  
[DAGA](#) 19.-22.03.2018 in München, **Paper Deadline 01.11.2017** [CfP]  
[ICPR](#) 20.-24.04.2018 in Beijing, China, **Paper Deadline 05.01.2018** [CfP]  
[Interspeech](#) 02.-06.09.2018 in Hyderabad, India, **Paper Deadline noch offen**

**SAVE THE DATE:** Die 13. ITG-Fachtagung Sprachkommunikation wird vom 10.-12. Oktober 2018 an der Universität Oldenburg stattfinden.

### Stellenangebote

- Die [EML European Media Laboratory GmbH](#) in Heidelberg sucht hochqualifizierte Mitarbeiter in den Bereichen Sprachtechnologie und Softwareentwicklung. Das EML entwickelt Software und Technologien in der automatischen Sprachverarbeitung, mit Schwerpunkt auf automatische Transkription, die in Telefonie und mobilen Anwendungen eingesetzt wird. Gegründet 1997 von Klaus Tschira, feiert das Unternehmen in diesem Jahr sein 20-jähriges Bestehen. [Link](#)