50 YEARS INSTITUTE OF ACOUSTICS AND SPEECH COMMUNICATION – 30 YEARS CONFERENCE ELECTRONIC SPEECH SIGNAL PROCESSING – 20 YEARS HISTORIC ACOUSTIC-PHONETIC COLLECTION

Rüdiger Hoffmann

Technische Universität Dresden, Institut für Akustik und Sprachkommunikation
ruediger.hoffmann@tu-dresden.de

Abstract: The speech researchers of the TU Dresden are pleased to welcome the participants of the 30th Conference on Electronic Speech Signal Processing (ESSV). Born in the time of the German re-unification, the conference series developed to a place of exchange between engineers, linguists, phoneticians and other people from academia and practice, who are interested in theory and application of speech technology. After several years, the conference returns to Dresden in 2019, where two further jubilees are coming up at the same year: the 50th anniversary of the foundation of the laboratory, which is now named Institute of Acoustics and Speech Communication, and the 20th anniversary of the foundation of the historic collection of this institute. Ten years ago, when all these numbers were correspondingly lower, the combination of the jubilees was appreciated by a "day of tradition" (September 23/24, 2009) and a special exhibition at the Technical Museum Dresden. The material of these events was collected in Volume 54 of the book series "Studientexte zur Sprachkommunikation" [5]. The present paper summarizes some selected parts of this material, which was published in German at that time, and gives a short update regarding the last decade.

1 30 Years Conference Electronic Speech Signal Processing (ESSV)

The first conference of the ESSV series took place in Berlin in 1990, the year of the German re-unification. This was not a simple coincidence. The specific significance of the event was already expressed by the invitation to join a *Gemeinsame Konferenz Elektronische Sprachsignalverarbeitung*. However, 30 years later, it seems to be useful to tell the today's participants why the word *gemeinsam* (common) was so essential.

Formally, the two German countries still existed when the conference was held. The organizer from the Western side was KLAUS FELLBAUM from the TU Berlin (West), which was the location where the meeting took place physically. The Eastern partners were WALTER TSCHESCHNER from the TU Dresden and DIETER MEHNERT from the Humboldt University Berlin (East) (Fig. 1). These partners maintained contact since some time before, making use of a workshop series in the GDR (the so-called *Klausurtagung Sprachkommunikation*, which reached a number of 10 meetings) and the BIGTech workshop in Western Berlin 1988. At the last *Klausurtagung* in Quedlinburg 1989, the process of re-unification was already visible, and the partners scheduled the "common" conference. By the way, the organization had to be splitted, due to the still existing different administrative systems.

The meeting was successful especially as a platform for discussions between Eastern and Western scientists, thus the organizers agreed to continue it in an annually rhythm. In the first years, the locations were always Dresden and Berlin, corresponding to the organizing institutions. In 1997, Cottbus was added as a standard place, because KLAUS FELLBAUM was

Figure 1. The founders of the ESSV. From the left: WALTER TSCHESCHNER (photograph by R. DIETZEL 1987), DIETER MEHNERT (photograph by R. DIETZEL 1992), and KLAUS FELLBAUM (private photograph 2005).

appointed Professor for Telecommunication at the TU Cottbus. To get more diversity, the conference moved to interesting other locations (Wolfenbüttel, Görlitz, Freiberg), but with the same organizers.

The 12th conference (Bonn 2001) was the first ESSV, which was hosted by another institute. This was the first trial to find a broader interest, which was continued in Karlsruhe 2003, Prague 2005 (the ESSV having the biggest number of papers ever), and Frankfurt am Main 2008. This successful way was consolidated in 2009, when the allocation of the conference was transferred to a newly founded association, called *Förderverein Elektronische Sprachsignalverarbeitung e. V.* Since then, the association decides annually, where the next conference will take place.

In some cases, the ESSV was combined with other events:

- 14th symposium *Mustererkennung* (pattern recognition) of the DAGM (Dresden 1992),
- 5th conference *Sprachkommunikation* (speech communication) of the ITG (Dresden 1998),
- 15th Czech-German Workshop "Speech Communication" (Prague 2005).

When the first conference was organized, nobody expected that the series would survive as long. There are certainly different reasons for this success. One aspect could be, that even in this period the big conferences like Eurospeech, ICSLP, and ICASSP have grown as much. Obviously, there is the need for an alternative type of local meetings, which combine a high scientific standard with a more intimate atmosphere. Following this approach, the ESSV has developed a unique, interdisciplinary profile over the years. The organizers are confidential, that the interest in the conference series will continue more years.

2 50 Years Institute of Acoustics and Speech Communication (IAS)

The institute which hosts the ESSV 2019 celebrates its 50th anniversary in just that year. Therefore, some remarks on the history of acoustics and speech communication at this place may be interesting.

The first acoustician of the (later) TU Dresden was AUGUST SEEBECK (1805–1849), who was the director of the *Technische Bildungsanstalt* from 1842 to 1849 [10]. The institution was upgraded to a *Polytechnikum* in 1871 and to a *Technische Hochschule* (TH) in 1890, where RICHARD HEGER (1846–1919) worked in the field of room acoustics [16]. But the biggest

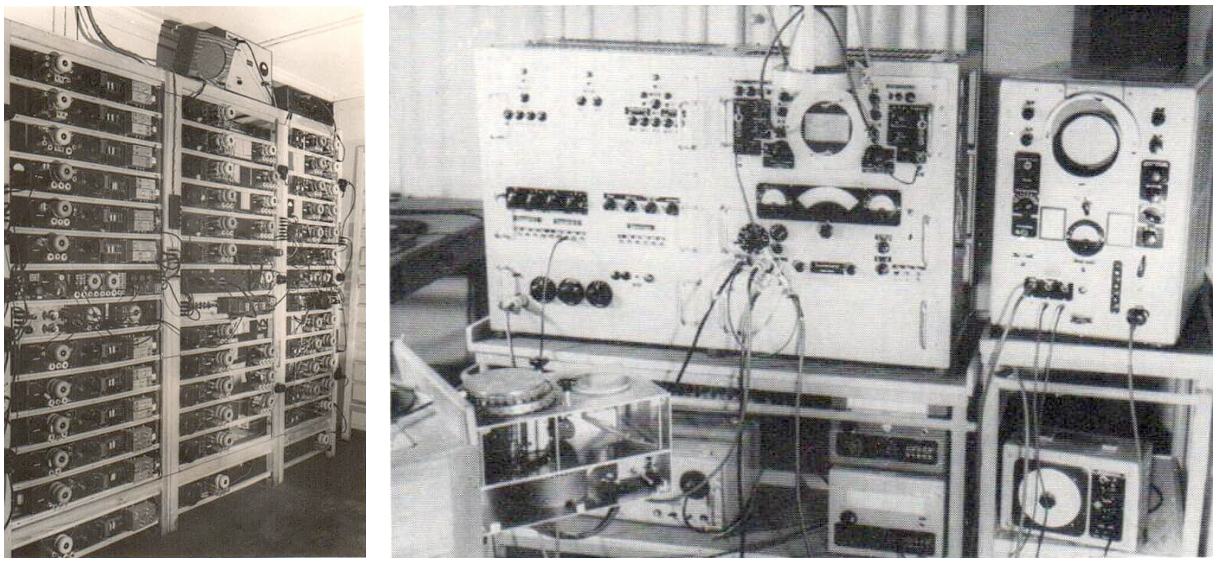

Figure 2. Early devices for speech research at the TH/TU Dresden. (a) The vocoder by EBERHARD KROCKER, 1958, (b) Speech analysis system using the heterodyne method (*Suchtonverfahren*) by WALTER TSCHESCHNER, 1958–1961. Photographs from the HAPS archive.

influence was due to HEINRICH BARKHAUSEN (1881–1956), who is widely known as the "father of the electron valve". He promoted the field of acoustics throughout Germany not only with the invention of the first applicable device for the measurement of loudness [1]. His aim to install a specific chair for acoustics in Dresden was prevented by Word War II.

When the TH Dresden (TU from 1961) was re-opened after the war destructions, an extra *Institut für Elektro- und Bauakustik* (electroacoustics and building acoustics) was established in 1950, renamed as *Institut für Technische Akustik* (technical acoustics) in 1967, directed by the renowned acoustician WALTER REICHARDT (1903–1985) [9]. Although speech processing was not in the main focus of this institute, a scholar of REICHARDT, S. STEINBACH, developed measuring systems for speech spectra. But the main interest in speech research was concentrated in the *Institut für Fernmeldetechnik* (telecommunication), which was founded by KURT FREITAG (1901–1977) in 1952.

Since the 1930s, the speech signal was investigated by engineers mainly with the focus on bandwidth reduction, resulting in the development of the vocoder, which took place more or less in parallel in Germany (K. O. SCHMIDT), the United States (H. DUDLEY), and the Soviet Union (V. A. KOTEL'NIKOV). Recognizing the importance of this development, FREITAG wanted to install a vocoder system at his Dresden institute, and his scholar E. KROCKER engineered a channel vocoder (Fig. 2 a) in the framework of his Dr.-Ing. thesis until 1957 [12].

This research was continued by WALTER TSCHESCHNER (1927–2004). He used the vocoder of the *Institut für Fernmeldetechnik* and up-to-date analysis methods (Fig. 2 b) for extensive investigations of the speech signal in the framework of his Dr.-Ing. (1961) and Habilitation (1968) theses. It was an enormous work, which is hard to image by people who are used to apply the advanced methods of computerized signal analysis. Importantly, TSCHESCHNER turned the focus from the telecommunication to the aspects of human-human and human-machine interaction, recognizing the significance of speech analysis and synthesis for the future development. He established contacts to international partners and to potential applicants in computer engineering, phonetics, rehabilitation, etc.

Among other reasons, the rapid development of science and engineering required deep changes in the academic landscape, which culminated in East and West around the year 1968. In the GDR, the so-called *3. Hochschulreform* (higher education reform) was organized by

the government. It was one of the main goals (and especially important for the engineering faculties), to establish a close cooperation between science, education and industry in the rapidly growing field of computer engineering. At the same time, the companies, which were already working in several fields of electronic data processing in the GDR, were merged to the *Kombinat Robotron* in 1969, which was a mighty partner for the universities henceforth.

When the Faculty of Electrical Engineering at the TU Dresden was re-organized in that sense, a scientific unit was built, consisting of the complete former *Institut für Technische Akustik*, the group of W. TSCHESCHNER from the *Institut für Fernmeldetechnik*, and a group working in electronic measurement from the *Institut für Hochfrequenztechnik* (radio engineering). It was the clear aim to form a laboratory having the power to address the future requirements, which were expected to occur at the interfaces of the computer with its environments as well as with the human user. This decision of the year 1969 was a very early one, compared to other universities, and we count it as the founding act of our recent institute. Its name, however, was still different and developed as follows:

1969: *Wissenschaftsbereich Kommunikation und Messwerterfassung* (Laboratory of Communication and Measuring), which included four working groups, originating from the aforementioned institutes.

1977: *Wissenschaftsbereich Akustik und Messtechnik* (Laboratory of Acoustics and Measurement), which was a simple re-naming to give the traditional field of acoustics more emphasis.

1990: *Institut für Technische Akustik* (Institute of Technical Acoustics), which was essentially a reminiscence at the name of the historic institute of W. REICHARDT after the political changes of 1989/90.

1999: *Institut für Akustik und Sprachkommunikation* (Institute of Acoustics and Speech Communication), considering the fact the institute had been reduced to these two chairs due to the stepwise personal reduction of the faculty.

Concentrating now to the speech communication group of the institute, the development was shaped by the following three university teachers:

- WALTER TSCHESCHNER, who started as a lecturer in the new laboratory, was appointed professor for Speech Communication in 1972. Until his retirement in 1992, he supervised 37 Dr.-Ing. and Habilitation theses in the field of speech synthesis and recognition. The scientific results were applied in numerous projects of the aforementioned *Kombinat Robotron* as well as the *Kombinat Musikinstrumente*. He developed numerous teaching activities in speech communication at the TU Dresden and outside. Although the Cold War strongly limited the interaction with the international community, the scientific level of his group remained high. Hence, the integration in the Western system posed no serious problems after the German re-unification.

- RÜDIGER HOFFMANN was appointed lecturer in Human-Machine Communication in 1986 and the successor at the chair of TSCHESCHNER in 1992. He guided the Dresden sub-project of the German *Verbmobil*, which offered a tremendous chance under the new political conditions in the 1990s. He supervised 36 Dr.-Ing. and Habilitation theses as well as numerous cooperation projects. He expanded the educational activities of the chair to the theory of signals and systems. This resulted in the re-naming of the chair to *Systemtheorie und Sprachtechnologie* (system theory and speech technology) in 2007

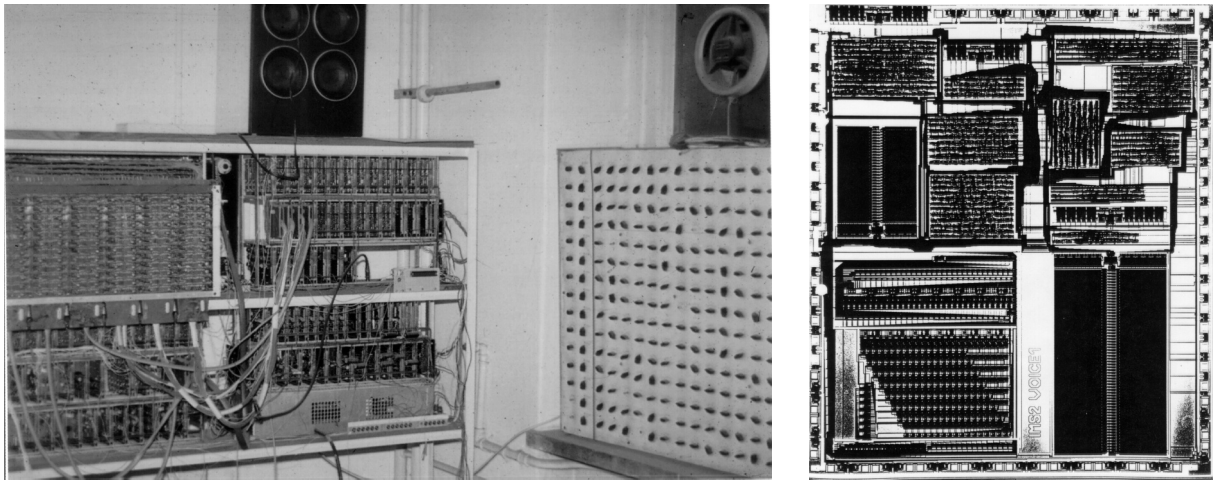

Figure 3. Examples for the rapid development of speech synthesis over merely two decades. Left: Partial view of the transistorized 3 formant speech synthesis system SYNI 2, TU Dresden 1974. Right: Formant speech synthesis integrated circuit VOICE 1, TU Dresden and Fraunhofer IMS2 Dresden 1993. Photographs from the HAPS archive.

and in the expansion of the scientific focus to cognitive systems (project UASR). Since his retirement in 2014, he has been responsible for the historical activities, which will be described in Section 3.

- PETER BIRKHOLZ started as Junior Professor for Cognitive Systems with tenure track as successor of HOFFMANN. His group researches at the confluence of speech technology, phonetics, and machine learning. The major topics are the multi-modal simulation of speech and language acquisition, the development of mechanical vocal tract models for embodied cognitive simulations of speech production, the investigation of Brain-Computer Interfaces to control speech synthesis systems and Silent-Speech Interfaces [15].

3 20 Years Historic Acoustic-phonetic Collection (HAPS)

The third jubilee refers to the historic acoustic-phonetic collection, which has its roots in the rather long history of the Dresden Speech Communication Group. A variety of old devices from the half century of research and development formed the base of the collection (Fig. 3).

If engineers are working in the field of speech technology, they need close contacts to phoneticians. In the case of the Dresden group, there was a traditional connection to the Phonetics Chair of the *Humboldt-Universität zu Berlin*, who was occupied in the later years by DIETER MEHNERT. This chair represented a long tradition, because Berlin was originally one of the most important places for the development of experimental phonetics in Germany. This tradition was materially visible in a collection of historic phonetic devices, which was compiled by D. MEHNERT. Unfortunately, his chair was closed due to the re-structuring of the Berlin universities in 1995. To save the material heritage, he proposed to merge his collection with its Dresden counterpart. The implementation of this idea, which was strongly supported by the custody, was finished in 1999, which we therefore count as founding year of the HAPS. The collection was thereby able to demonstrate the development of speech technology from the early years of experimental phonetics until the introduction of the computer in the laboratories of the phoneticians and engineers.

The next essential step in the development of the HAPS followed in 2005, when the Institute of Phonetics of the Hamburg University was closed. This institute originated from the Pho-

Figure 4. Since 2015, the objects of the HAPS are shown in three display rooms in wing C of the Barkhausen building, which are devoted to the thematic fields of experimental phonetics, speech production and synthesis (photograph), and speech analysis and recognition.

netic Laboratory of the Hamburg Colonial Institute, which was directed since 1910 by GIULIO PANCONCELLI-CALZIA (1878–1966), who was one of the most important scientists in the early years of experimental phonetics. His laboratory survived the war destruction of Hamburg and was available nearly in the original state, when the institute was closed. The equipment was transferred to the HAPS based on an agreement of the deans of the involved departments. The Faculty of Electrical Engineering and Information Technology supported the installation generously by the provision of rooms (Fig. 4) and other infrastructure.

The availability of this representative collection offers the chance to investigate numerous research themes. The work, which was done until now, may be subdivided in three main parts:

- **Cataloguing and publishing the objects.** In many cases, it was not easy to evaluate the history and the application of the old objects. Numerous detailed investigations have been performed and were published mainly at conferences. The publication of a catalogue in a printed and an electronic version is the main goal, which is hitherto fulfilled partially: The historic phonetic devices are described in a first volume, which was published by D. MEHNERT (text) and R. DIETZEL (photographs) in 2012 [13]. A complementary volume on the people behind the exhibits appeared in 2016 [14]. The second volume of the catalogue, entitled "Historic devices of speech acoustics", is expected to be completed in 2020. The planned online availability will be provided using the database DAPHNE, which is used by the State Art Collections in Dresden and will be adopted by the custody of the TU Dresden, too.

- **Kessel's mechanical voices.** When the Hamburg collection came to Dresden, a set of small mechanical voices, pronouncing vowels and short words, excited special interest. The investigation of these objects developed some own dynamics in two directions. Because it was known that JOHANNES KESSEL (1839–1907), a professor for otology in Jena, proposed to use the voices for training purposes with hard hearing patients in 1899, a biography of KESSEL was elaborated in cooperation with an expert in oto-rhino-laryngology and its history [7]. Secondly, it turned out that the voices came originally

Figure 5. Examples for the development of mechanical models of the articulation tract. Photographs TU Dresden / HAPS.
Left: Resonators for the sounds *a* and *e* by C. KRATZENSTEIN 1781 [11], designed with trial and error. Replicas by C. KORPIUN 2005.
Middle: Resonators for the sounds *a* and *e* by T. CHIBA and M. KAJIYAMA 1941 [2], designed basing on X-ray data. 3D-printed replicas by T. ARAI 2017.
Right: Resonators for the sounds *i* and *o* by P. BIRKHOLZ and collaborators [4], designed basing on Magnetic Resonance Imaging (MRI) and 3D scanning data. 3D-prints from 2018.

from the toys industry, where they appeared as first industrial spin-offs of the speaking machine of W. V. KEMPELEN [6]. A number of gifts and purchases enabled the extension of this part of the HAPS, demonstrating the development of voices for toys and books over two centuries.

- **BMBF project *Sprechmaschine* (speaking machine).** The German Science Council published a recommendation on university collections in 2011, stressing their importance for teaching and research [3]. Since then, university collections benefit from more public support. The TU Dresden succeeded in 2017 with a joint application at the German Ministry of Education and Research (BMBF). This project, called *Sprechmaschine* (speaking machine), draws a long bow from the early attempts of mechanic speech synthesis up to the state-of-the-art research in articulatory modeling (Fig. 5) [8]. In this project, the group of P. BIRKHOLZ (project manager) cooperates with the chairs of applied linguistics, media design, and industrial design engineering. The State Art Collections Dresden support the project with experience from its historic mathematic-physical collection (*Mathematisch-Physikalischer Salon*).

References

[1] BARKHAUSEN, H.: *Ein neuer Schallmesser für die Praxis*. In: *Zeitschr. f. technische Physik* 7 (1926) 12, pp. 599–601.

[2] CHIBA, T.; KAJIYAMA, M.: *The Vowel: Its Nature and Structure*. Tokyo-Kaiseikan Pub. Co. 1941.

[3] *Empfehlungen zu wissenschaftlichen Sammlungen als Forschungsinfrastrukturen*. In: *Empfehlungen zu Forschungsinfrastrukturen.* Köln: Wissenschaftsrat 2011, pp. 209–282.

[4] FLEISCHER, M.; MAINKA, A.; KÜRBIS, S.; BIRKHOLZ, P.: *How to precisely measure the volume velocity transfer function of physical vocal tract models by external excitation.* PLoS ONE 13 (2018) 3, e0193708. doi: 10.1371/journal.pone.0193708

[5] HOFFMANN, R. (Hrsg.): *Elektronische Sprachsignalverarbeitung 2009, Band 2: Tagungsband des Traditionstages, Dresden, 23./24. 9. 2009; Katalog der Ausstellung SprachSignale, Dresden, 23. 9. 2009 bis 21. 3. 2010*. Dresden: TUDpress 2010 (Studientexte zur Sprachkommunikation; 54).

[6] HOFFMANN, R.: *Voices for toys – First commercial spin-offs in speech synthesis*. In: HOFFMANN, R.; TROUVAIN, J. (eds.): *HSCR 2015. Proceedings of the First International Workshop on the History of Speech Communication Research, Dresden, September 4-5, 2015.* Dresden: TUDpress 2015 (Studientexte zur Sprachkommunikation; 79), pp. 60–70.

[7] HOFFMANN, R.; LÖBE, L.-P.; PFEIFFER, W.: *"Ich holte meine Prager Schriften". Leben und Werk des Otologen Johannes Kessel (1839–1907)*. Dresden: TUDpress, 2015 (Studientexte zur Sprachkommunikation; 80).

[8] HOFFMANN, R.; BIRKHOLZ, P.; GABRIEL, F.; JÄCKEL, R.: *From Kratzenstein to the Soviet Vocoder: Some Results of a Historic Research Project in Speech Technology.* In: KARPOV, A.; JOKISCH, O.: POTAPOVA, R. (eds.): *Speech and Computer. 20th International Conference SPECOM 2018 Leipzig, Germany, September 18-22, 2018, Proceedings*. Cham: Springer Nature 2018 (Lecture Notes in Artificial Intelligence; 11096), pp. 215–225.

[9] KÖLTZSCH, P. (Hrsg.): *Festschrift zum Ehrenkolloquium Reichardt – Kraak – Wöhle, 4. Juli 2003. Zur Entwicklung des Fachgebietes Technische Akustik und des akustischen Instituts an der TH / TU Dresden in den Zeitläuften des 20. Jahrhunderts*. TU Dresden 2003; 2., erweiterte Aufl. 2003.

[10] KÖLTZSCH, P.: *August Seebeck – Der erste Akustiker der Dresdner "Universtät"*. In: MEHNERT, D.; KORDON, U.; WOLFF, M. (Hrsg.): *Systemtheorie, Signalverarbeitung, Sprachtechnologie. Rüdiger Hoffmann zum 65. Geburtstag*. Dresden: TUDpress 2013 (Studientexte zur Sprachkommunikation; 68), pp. 295–302.

[11] KRATZENSTEIN, C. G.: *Tentamen resolvendi problema, Petersburg 1781*. Übersetzt und kommentiert von CHRISTIAN KORPIUN. Dresden: TUDpress 2016 (Studientexte zur Sprachkommunikation; 82).

[12] KROCKER, E.: *Aufbau und Untersuchung eines Übertragungssystems für synthetische Sprache*. In: Wissenschaftl. Ztschr. der TH Dresden 6 (1956/57) 4, pp. 757–776.

[13] MEHNERT, D.: *Historische phonetische Geräte. Katalog der historischen akustisch-phonetischen Sammlung (HAPS) der Technischen Universität Dresden, erster Teil*. Mit 393 Fotografien von ROLF DIETZEL. Dresden: TUDpress 2012 (Studientexte zur Sprachkommunikation; 62).

[14] MEHNERT, D.; PÉTURSSON, M.; HOFFMANN, R.: *Experimentalphonetik in Europa. Institute, Wissenschaftler, Leistungen Ende des 19. / Anfang des 20. Jahrhunderts*. Dresden: TUDpress 2016 (Studientexte zur Sprachkommunikation; 85).

[15] *TU Dresden, website Juniorprofessur Kognitive Systeme*. URL (15. 1. 2019): `https://tu-dresden.de/ing/elektrotechnik/ias/juniorprofessur-fuer-kognitive-systeme/die-professur?set_language=en`

[16] VOSS, W.: *Dresdens große Mathematiker.* In: *Dresdner Universitätsjournal*, Sonderausgabe 2001-1, pp. 18–19.