Professor Dr.-Ing. Hermann Ney
Human Language Technology and Pattern Recognition
RWTH Aachen University, D-52056 Aachen, Germany
21-Feb-2021

EDUCATION:

1977	Diplom	Physics	University of Göttingen, Germany
1982	Dr.-Ing.	Electrical Engineering	University of Braunschweig, Germany

PROFESSIONAL POSITIONS:

1977 – 1993	**Philips Research Laboratories,** Hamburg and Aachen, Germany; 1984-1999 head of research department *pattern recognition, speech recognition and understanding*;
1993 – today	**RWTH Aachen University**, Aachen, Germany; professor of computer science.

EXTENDED VISITS:

1988	visiting scientist, *Speech Research Department*, **AT&T Bell Labs,** Murray Hill, NJ;
1997	visiting professor, **ICSI**, Berkeley, CA;
2002-2021	visiting professor, research team *Pattern Recognition and Human Language Technology*, **Universidad Polytecnica de Valencia**, Valencia;
2011-2014	DIGITEO professor, senior chair at **LIMSI-CNRS**, Paris.

RESEARCH AREAS:

- machine learning and data-driven approaches to speech and language processing;
- speech recognition and machine translation of written and spoken language;
- handwriting recognition, sign language recognition and translation.

PUBLICATIONS:

- more than 800 papers in international conferences and journals;
- many invited talks and best paper awards;
- h-index of more than 100 and about 60000 citations (based on *Google Scholar*; more than 25000 citations in 2016-2020).

AWARDS:

- fellow: ISCA (2009), IEEE (2011) and ACL (2020);
- distinguished lecturer: ISCA (2012/3) and IEEE (2016/7);
- 2005 IEEE Signal Processing Society: Technical Achievement Award:
  ... for contributions to the advancement of the theory and performance of speech and language technology, including language modeling, search algorithms and machine translation;
- 2013 IAMT Award of Honour (IAMT = Int. Association of Machine Translation):
  ... for his lifelong contributions to the field of machine translation, especially in the area of statistical machine translation, and for his leadership of the influential Human Language Technology and Pattern Recognition Group of the Computer Science Department at RWTH Aachen;
- 2016 Advanced Grant (topic: *sequence processing for speech recognition and machine translation*) of the European Research Council (ERC);
- 2019 IEEE James L. Flanagan Speech and Audio Processing Award:
  ... for pioneering contributions to statistical and computational modeling for speech recognition and machine translation.
- 2021 ISCA Medal for Scientific Achievement:
  ... for pioneering and seminal contributions to data-driven methods for automatic speech recognition and machine translation.

Professor Dr.-Ing. Hermann Ney
Human Language Technology and Pattern Recognition
RWTH Aachen University, D-52056 Aachen, Germany
21-Feb-2021

SELECTED CONTRIBUTIONS TO SCIENCE AND TECHNOLOGY:

- 1990 *dynamic programming beam search* for large-vocabulary continuous speech recognition:
  based on dynamic construction of the search space using a lexical prefix tree, which allowed real-time recognition on standard hardware at that time and which is still used today in form of weighted finite state transducers.
- 1993 *Philips Vienna:* commercial product for text dictation;
  first product for large-vocabulary continuous speech recognition.
- 1995 *Kneser-Ney smoothing* for language modelling:
  until 2012, this was the best performing language model and is still widely used today.
- 1997 *Philips Dialogue Systems*: first commercial deployments
  of spoken dialogue systems over the telephone (based on Harald Aust's PhD Thesis: *Speech Understanding and Dialog Modelling in Natural Language Information Systems*, RWTH, July 10, 1998).
- 1999 data-driven *phrase based machine translation*:
  introduction of bilingual word groups as basic units rather than single words as in IBM's approach; this approach was later called *phrase based approach* and led to a revolution in the field.
- 2001 *public toolkit GIZA++* for word alignment in machine translation:
  this toolkit lowered the entrance barrier for new teams and is still widely used.
- 2005 *speech translation* in EU project TC-STAR:
  first research protoype for *unlimited-domain* translation of spoken language for *real-life data* like speeches given in the EU parliament.
- 2008 *Google's translation service*:
  Franz Och and more former PhD students from RWTH joined Google research and utilized the *phrase based approach* for this service.
- since 2012 *deep neural networks* for speech and language technology:
  – LSTM-RNN for language modelling (Sundermeyer et al., Interspeech 2012):
  until today, this is the best method for language modelling.
  – DNN-based feature extraction from speech waveform
  (Tueske et al., Interspeech 2014).
  – acoustic modelling using various DNN types (Luescher et al., Interspeech 2019):
  best recognition results on *Switchboard* and *Librispeech*.
  – DNN-based hidden Markov models for machine translation
  (Sundermeyer et al., EMNLP 2014; Wang et al., ACL 2018).

SYSTEMS, PROJECTS AND FUNDING:

- successful participation of RWTH systems in international evaluation campaigns for speech recognition, handwriting recognition and language translation;
- EU and French funding for various projects like TC-STAR, QUAERO, EU-BRIDGE and QT21;
- American funding (DARPA and IARPA) in projects like GALE, BOLT and BABEL;
- funding by companies like Google, Nuance Communications, Ebay and Ford.

Professor Dr.-Ing. Hermann Ney
Scientific Achievements in Detail
21-Feb-2021

- **Automatic Speech Recognition (ASR):**
  Since his PhD work in 1977, Hermann Ney has been working on speech recognition. In 1993 when he joined RWTH University, he continued this research area. His research topics cover virtually all aspects of automatic speech recognition and have resulted in various scientific contributions:

  1. efficient search methods (dynamic generation of the search space by using history dependent lexical prefix trees and look-ahead based on language and acoustic models) **2. methods for language modelling and smoothing**, which are today known as *Kneser-Ney* modeling and used in many systems; **3. confidence measures**, based on word posterior probabilities; **4. discriminative training procedures** for reducing the error rate; **5. artificial neural networks** that are used at levels of acoustic, phonetic and language modelling.

  Many of today's system for ASR make use of these concepts in various forms. The ASR software (name: RASR and RETURNN for neural nets in ASR) of the chair is available as a free public toolkit for academic research.

- **Machine Translation (MT):**
  Hermann Ney has been working on statistical machine translation since 1995. At that time, after IBM had stopped working on MT, there were only a few teams worldwide that considered the statistical approach to be promising for MT and had active research activities. Hermann Ney and his team introduced various concepts in statistical MT. Many of them have been adapted by other teams and are now widely used in many research systems and operational systems:

  1. various alignment models and their efficient implementation, which resulted in the 'public toolkit' GIZA++; **2. various search variants** for sentence generation, that are based on dynamic programming and *beam search*; **3. bilingual phrases and alignment templates**, that better capture the context of the translation process than the single-word IBM models; **4. log-linear model combinations**, by which various types of dependencies can be combined; **5. morpho-syntactic preprocessing** and its combination with statistical MT; **6. system combination**, by which the outputs of several MT engines can be combined; **7. confidence measures**, by which the reliability of tranlsations can be quantified; **8. direct Hidden Markov models**, which are based on deep neural networks.

  In total, these methods have resulted in a revolutionary development of the MT field. The MT software of the chair (name: JANE for phrase-based MT and RETURNN for neural based MT) is available as a free public toolkit for academic research.

- **other research areas:**
  In addition to speech recognition and machine translation, Hermann Ney has also worked on related research areas like sign language recognition and translation, handwriting recognition, image and object recognition.

- **PhD students:**
  Up to now, there have been 63 students in Hermann Ney's team who have received the PhD degree. Many of them joined international research teams such as Amazon, Google, Ebay, IBM and Nuance Communications. Franz Och, who was one of Hermann Ney's first PhD students on statistical machine translation, joined Google and introduced the statistical machine translation to Google. With more of Hermann Ney's PhD students joining him, he introduced the service *Google Translate* in 2008.

Professor Dr.-Ing. Hermann Ney
Major Research Projects (2004-2020)
21-Feb-2021

During the last 15 years, H. Ney was involved in a number of European and American projects on human language technology, including speech recognition, machine translation and text image recognition. Among these projects, there were also projects on a very large scale, like TC-STAR, GALE and QUAERO. In these projects, H. Ney had a leading role as the director of the research track on machine translation, which involved managing 10–25 scientists across several countries. Here is a summary:

- TC-STAR, integrated project funded by EU, 2004-2007:
  The goal of this project was the recognition and translation of speeches given in the European Parliament. The languages considered were Spanish and English. This was the *first* project that developed operational research prototype systems for *speech translation of real-life data in an unlimited domain*.
- GALE and BOLT, funded by US DARPA, 2006-2015:
  GALE was the largest project that was ever carried out in history on speech and language; it comprised 200 million USD in total. In addition to the American research teams (about 25), there were only about 5 non-American research teams, among them RWTH. The overall goal of GALE and BOLT was to build high performance systems for speech recognition and machine translation. For speech, the task was to recognize broadcast news and broadcast conversations in Arabic and Mandarin and to translate them into English. For written language, the task was the translation of newswire articles and texts from newsgroups in the internet. BOLT was a follow-up project that focused on more informal language.
- QUAERO, funded by French government, 2008-2013:
  In this project, the author and his team worked on speech recognition, machine translation and image text recognition, in particular for European languages.
- additional projects, funded by EU 2012-2017:
  In these projects (TransLectures, EU-BRIDGE and QT21), the author's team worked on speech recognition and machine translation for acadamic lectures and public lectures like TED.
- SEQCLAS: funded by European Research Council (ERC), 2016-2021:
  basic research for speech recognition and machine translation.
- SIGNSPEAK, funded by EU 2009-2012:
  The topic was the recognition and translation of sign language (weather forecasts on German TV). The author's team developed the first-prototype system for recognizing and translating *real-life* sign language data.
- operational prototype systems in research:
  The author's team has developed speech recognition for many (10+) languages and machine translation systems for many (10+) language pairs. These systems participated in many international evaluation campaigns and had leading positions in most of them.

Public Software Packages

H. Ney's team is one of the very few academic teams in the world that has its own proprietary software for high-performance HLT systems. In particular, there are fully fledged systems for speech recognition, machine translation and handwriting recognition that have been proven to be very competitive in many evaluation campaigns. The source code of the software is public and can be used free of charge by academic research; for commercial use, there is a specific license. Here is a list of the major public software toolkits:

- RASR: a fully fledged system for speech recognition.
- JANE: a fully fledged system for statistical machine translation.
- RWTH OCR: a fully fledged system for text image recognition.
- RWTH LM: a toolkit for feedforward and recurrent neural network language modeling.
- RETURNN: a toolkit for training deep neural network for speech recognition and language translation.
- GIZA++: training of statistical word alignments for machine translation.
- Sequitur G2P: a trainable grapheme-to-phoneme converter.