

Bridging the Language Divide

Alex Waibel and the InterACT Team

Carnegie Mellon University

Karlsruhe Institute of Technology

Mobile Technologies, LLC

alex@waibel.com

- Dilemma:
 - Living in the Global Village
 - Globalization, Global Markets
 - Increased Exchange and Communication
 - European/International Integration
 - Cultural Diversity:
 - Beauty, Identity, Language, Culture, Customs
 - Pride and Individualism
 - Language Ability
 - Challenge:
 - Providing Access to Global Markets and Opportunities
↔ Maintaining Cultural Diversity/Individuality

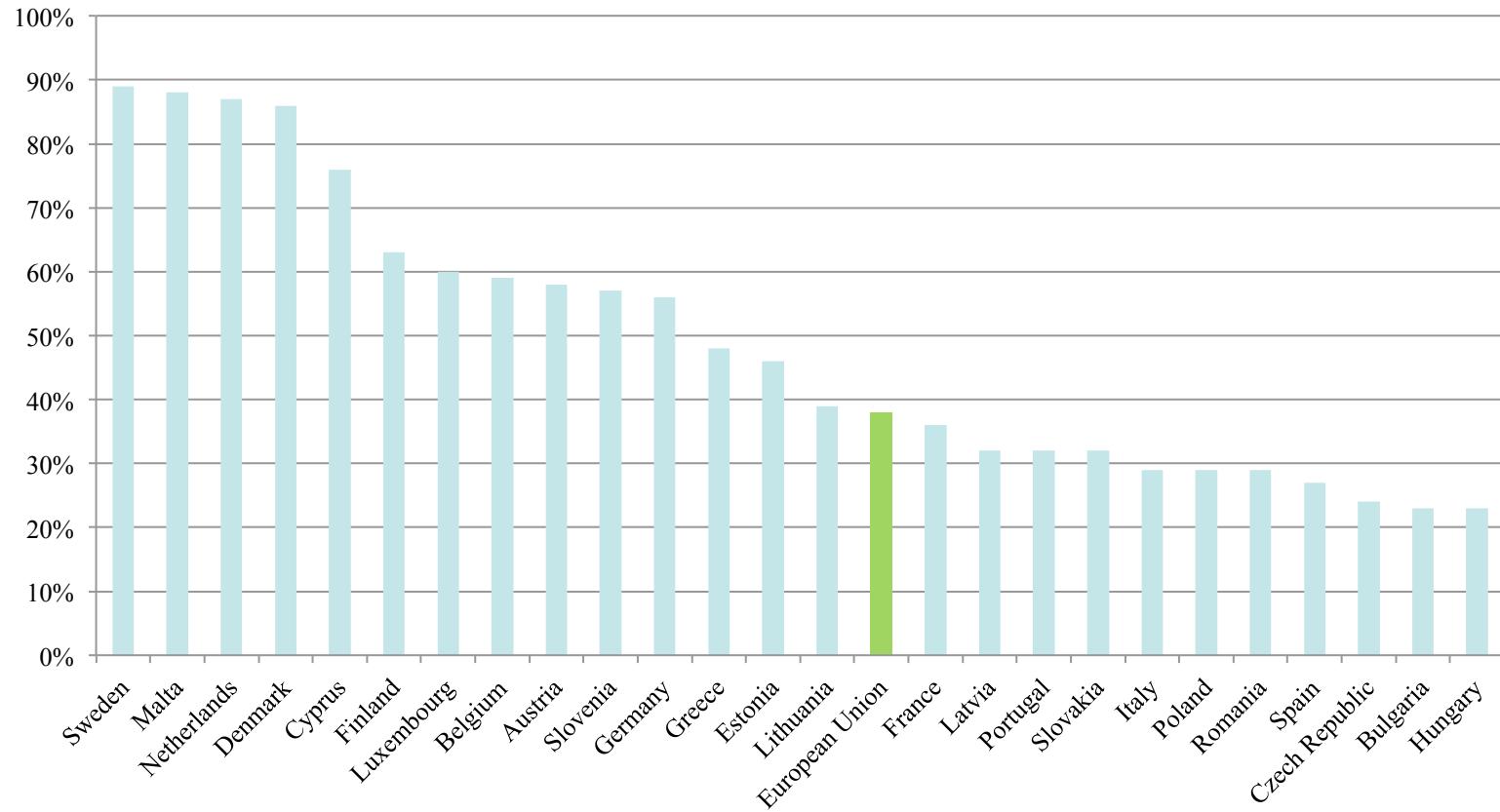
The Magnitude of the Problem

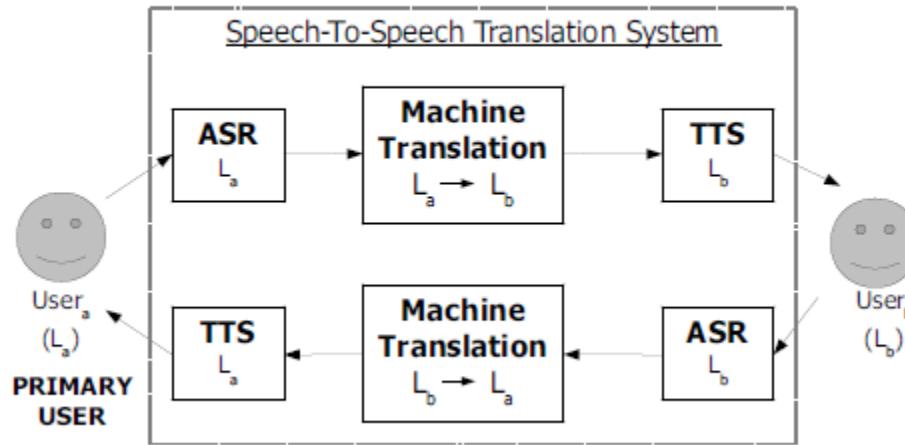
- Today Almost all Translation Done by Human Effort (>99%)
- 400,000 translators worldwide. 150,000 in Europe
- \$5.7 Billion dollar industry in 2006.
- European Union: 1.3 Billion Spent on Translation
 - 506 Language Directions to Translate
 - Current Effort Insufficient to Keep up with Needs of 27 Member States
- Worldwide 6000 Languages
 - 36,000,000 language directions !!!
- Actual translation work is currently only about 10% of translatable text.
- Translation needs are growing 25% -35% per year
- And that's just for Text.....

- Conferences
 - Estimated 300,000 conference per Year in Europe
 - Few Professional Simultaneous Translators
 - 1% or Less are Translated
- Internet
 - On You-Tube, Every Minute 13 Hours of New Videos
- Television
 - Satellite, Cable: Virtually Unlimited Channels
- Lectures
 - Government, Universities, Corporations
- Meetings
- Telephone Conversations
- Travel Dialogs/Encounters

In Europe:

English language knowledge (not mother tongue)

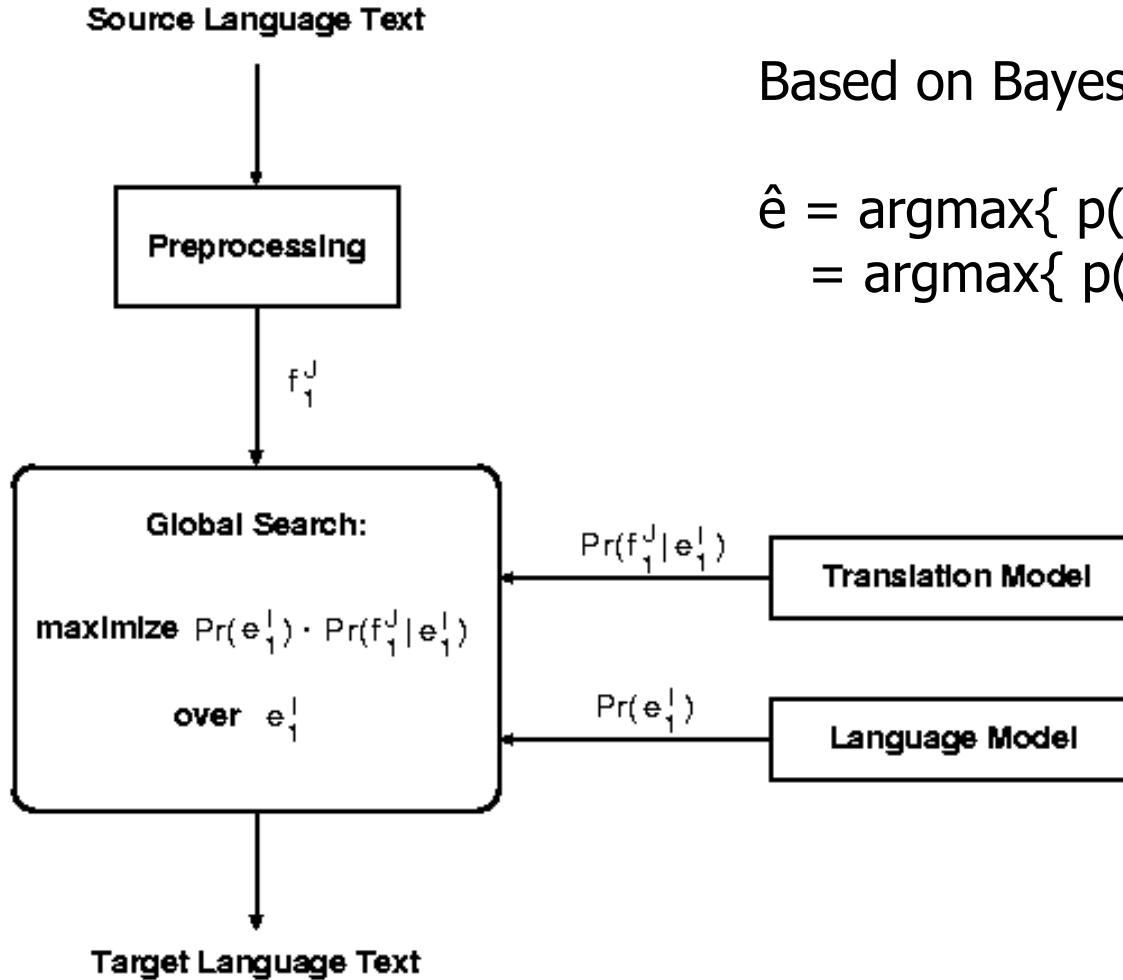




To Build a Speech Translator for a New Language

- 6 Component-Engines
- Each is in Principle Language Independent,
but Requires Language Dependent Parameters/Models
- Models are Automatically Trained but Require Large Corpora
- Certain Language Dependent Peculiarities Exist
- Performance and Human Factors are Important Issues

Statistical Translation Approach



Based on Bayes' Decision Rule:

$$\begin{aligned}\hat{e} &= \operatorname{argmax}\{ p(e | f) \} \\ &= \operatorname{argmax}\{ p(e) p(f | e) \}\end{aligned}$$

Progression of Technologies:

- Domain Limited, Clear Speaking Style (late 80's-91)
 - Janus (first European&US speech-to-speech system)
 - ATT, NEC, ATR
- Domain Limited, Spontaneous ('91-'00)
 - Janus II/III (work on 20 languages),
Verbmobil, Nespole, Enthusiast,
C-STAR, ATR, ETRI, NLPR,...
- Fieldable, Maintainable, Spontaneous
 - Transtac, Babylon, Phraselator, Jibbigo, U-STAR
- Domain Unlimited Speech Translation
 - Parliamentary Speeches (TC-STAR)
 - Broadcast News (GALE)
 - Lectures, Seminars (InterACT, STAR-DUST, TC-STAR)

Domain Limited Consecutive Translation

Technologies for Cross-Lingual Dialog

iPhone sync is complete.

OK to disconnect.

jibbigo



App Store > Travel > Jibbigo Speech Translator English Spanish

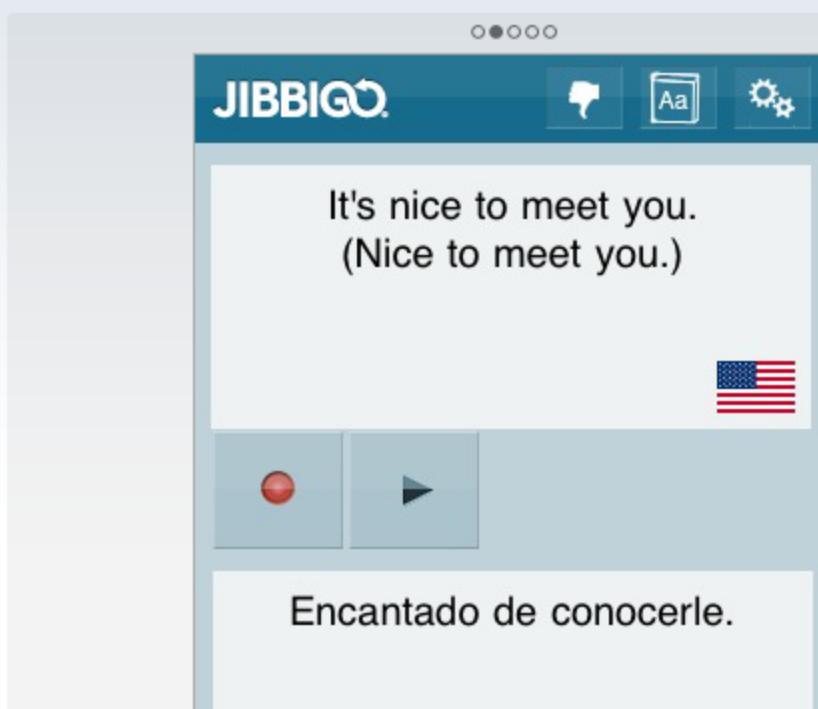
Jibbigo LLC**Jibbigo Speech Translator English Spanish**

Category: Travel
 Released Oct 21, 2009
 Seller: Jibbigo LLC
 © 2009 Mobile Technologies LLC
 Version: 1.0
 158 MB

\$27.99

BUY APP

Rated 4+

**APPLICATION DESCRIPTION**

Jibbigo is a bi-directional, natural speech-to-speech translation app that lets you converse with a speaker of another language by naturally spoken sentences.

Jibbigo is not a dictionary and not a phrase book, but a speech translator: You simply speak a sentence in English or Spanish into Jibbigo, and it speaks the sentence aloud in the other language, much like a personal human interpreter would. Jibbigo also shows the recognition and translation in English and Spanish as text on the app screen, so you can be sure your translation is accurate to what you said.

- ## Jibbigo:
- Real-Time Translation
 - Spanish - English
 - Japanese, Chinese
 - Arabic, ...
 - 40,000 Words
 - No Server Necessary



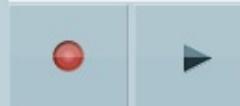
JIBBIGO.



Hello. Nice to meet you!
(Hello, nice to meet you.)



Hola, encantado de conocerle.







New year. New resolutions. New apps.

Weight Watchers

Free • Resolved to stay fit? Now you can get healthy tips and free recipes from one of the top names in weight loss – all year long. And if you're a Weight Watchers member, you can also track and calculate your points.

e-Secure

Free • Manage your Protection One home alarm system wirelessly from your iPhone. Turn your alarm on or off, receive text alerts, and more – even when you're away from the house.

RedEye

Free • Too many remotes cluttering your home? Turn your iPhone into a handy universal remote that can control your DVR, DVD player and stereo through your RedEye system.

Green Outlet

\$2.99 • Run a more eco-friendly home in 2010. Just enter in what appliances you use and how often each day and Green Outlet will help you identify where you can save energy – and money.

Master Control

Free • Looking to improve your soccer skills this year? This elite Nike training app features instructional videos from FC Barcelona coaches. Get tips on ball control, see drills the pros use and even learn some of their signature moves.

Dragon Dictation

Free • Send messages easier and faster, without typing a single word. Dragon Dictation recognizes and transcribes what you say, and lets you paste it directly into an email or text.

Yelp Monocle

Free • Find a new place to eat in a whole new way. Just point your iPhone camera down any street and Yelp Monocle will overlay restaurant ratings, pricing info and even how many miles away they are, right on your screen.

Ustream

Free • Use your iPhone to stream live video to your family and friends, as you shoot it. Through your Ustream channel, you can also share your videos on Twitter, Facebook and YouTube.

Jibbigo

\$24.99 • Carry a personal Spanish translator in your pocket. Just tap to record any phrase and Jibbigo will play it back in Spanish – or any Spanish phrase back into English.

Viper SmartStart

Free • Turn your iPhone into the ultimate keyless remote. With SmartStart installed in your vehicle, all you need is your iPhone to lock, unlock, or even start your car from just about anywhere.

ecobee

Free • Control your home's ecobee Smart Thermostat even when you're away. Adjust the temperature right from your iPhone so you can make sure you conserve energy costs but still come back to a cozy house.

New York Subway

\$0.99 • There's an amazing new way to find the nearest New York subway station. Just aim your iPhone's camera in any direction and NY Subway will superimpose station information and distances right on your screen.

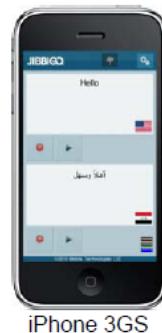


It's a new year with new amazing apps. From live streaming video to augmented reality, there are over 100,000 apps for just about anything. Only on the iPhone and the nation's fastest 3G network.

at&t | iPhone 3GS

- iTunes & Android App Stores:
 - English, Spanish, French, German, Japanese, Chinese, Korean, Filipino Iraqi
- Cost:
 - Free Online Version Jibbigo.net
 - Off-Line: No Network Needed \$4.99
- Outside of App Store:
 - Thai, Pashto, Dari,
 - Other Languages in Preparation

Supported Devices



iPhone 3GS



iPod Touch



iPad

iPhone-OS Compatible Devices
(First systems commercially deployed in Oct 2009)



Nexus One



Motorola Droid

HTC Droid
IncredibleSamsung i9000
Galaxy

Android-OS Compatible Devices

Humanitarian Deployment



Communication

- How it is Done:
 - Human Interpreters
 - Charts, Dictionaries
- Limitations/Problems:
 - Limited Supply!!
 - Fidelity/Trust/Security
 - Number of Languages



Unlimited Domain Simultaneous Speech Translation Technologies

Domain Unlimited Translators for:

- TV/Radio Broadcast Translation
- Translation of Lectures and Speeches
- Parliamentary Speeches (UN, EU,..)
- Telephone Conversations
- Meeting Translation



你们的评估准则是什么

Language Barriers



Studentin aus Tunesien

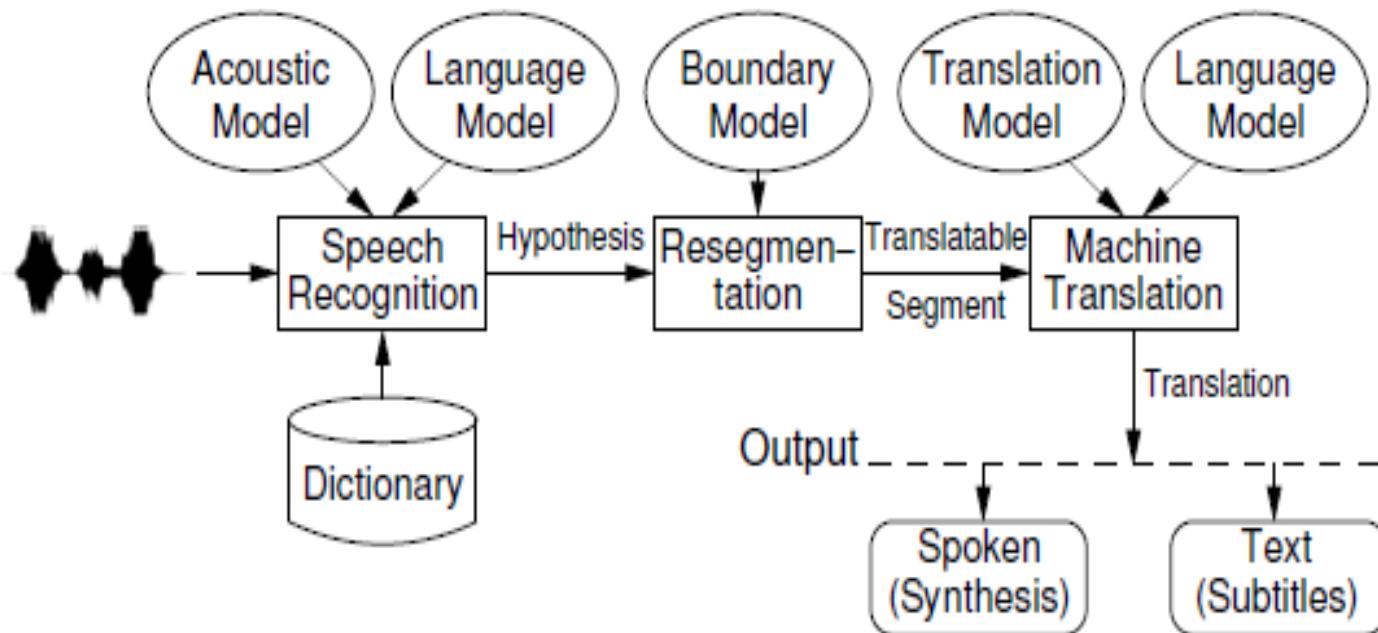


die Vorlesungen waren unverständlich

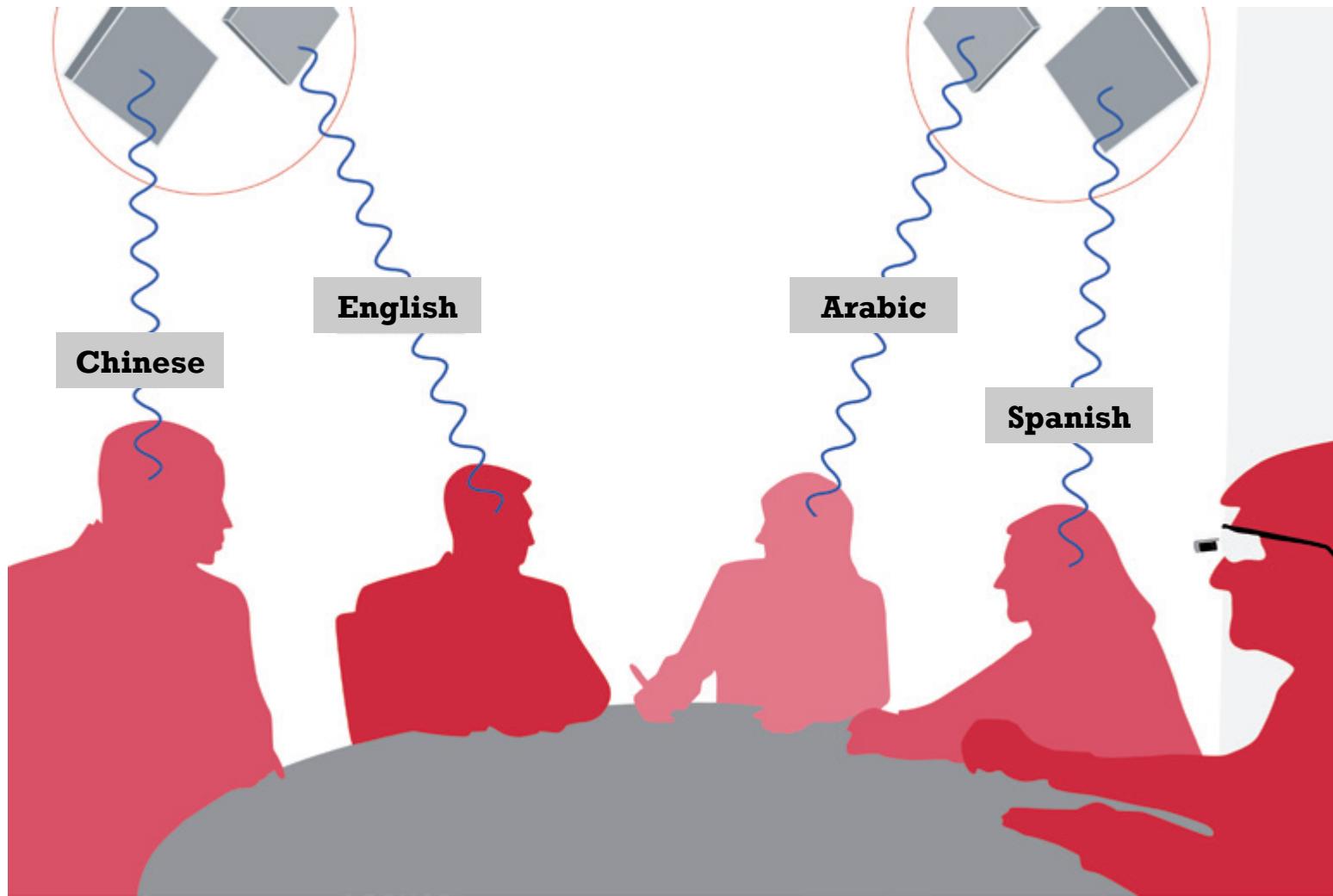
A close-up photograph of a young woman with long, dark brown hair. She is looking directly at the camera with a neutral expression. The background is blurred, showing what appears to be an indoor setting with warm lighting.

Translation of Speeches









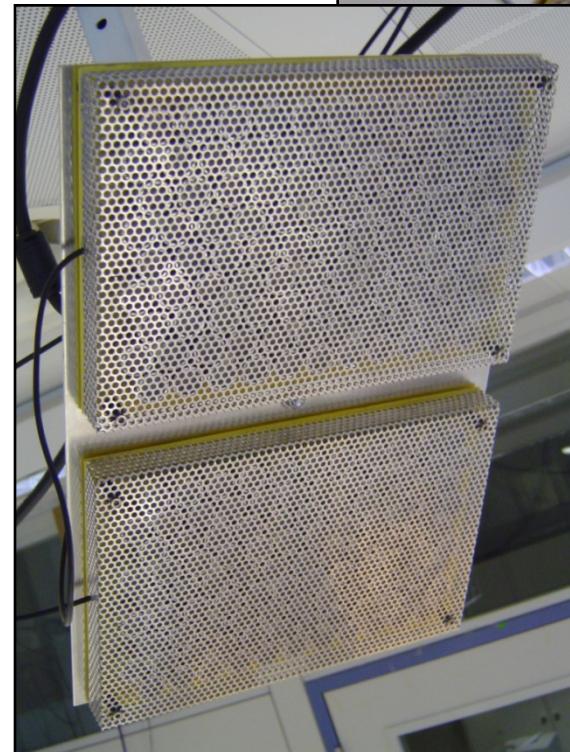
3. ---Seeing Personal Translations



- Technology: Heads-up Display Goggles
 - Create Translation Goggles
 - Run Real-Time Simultaneous Translation of Speech
 - Text is Projected into Field of View of Listener
 - Translations are Seen as Text Captions Under Speaker
 - Output: Spanish, German,...

Hearing Personal Translations

- Technology: Targeted Audio
 - Research under EC Project CHIL
(Build Inobtrusive Computer Services)
 - Project Partner, Daimler-Chrysler
 - Array of Ultra-Sound Speakers
- Result: Narrow Sound Beam
 - Audible by one Individual Only
 - Others not Disturbed
 - Multiple Arrays Could Provide Multiple Languages
 - Steerable
 - Recognize/Track Individual Listener and Keep Language Beam on Target



- English->Spanish Lectures
First Research-Prototype CMU 2005
- German Lectures, KIT '10
- Cloud Based Services, MobileTech, '10
- Transition to a Lecture Service
 - First Beginning: 4 Lectures, 2012
 - EU-BRIDGE



Shanghai Lecture, Simultaneous Translation Service

<http://lt.anthropomatik.kit.edu>

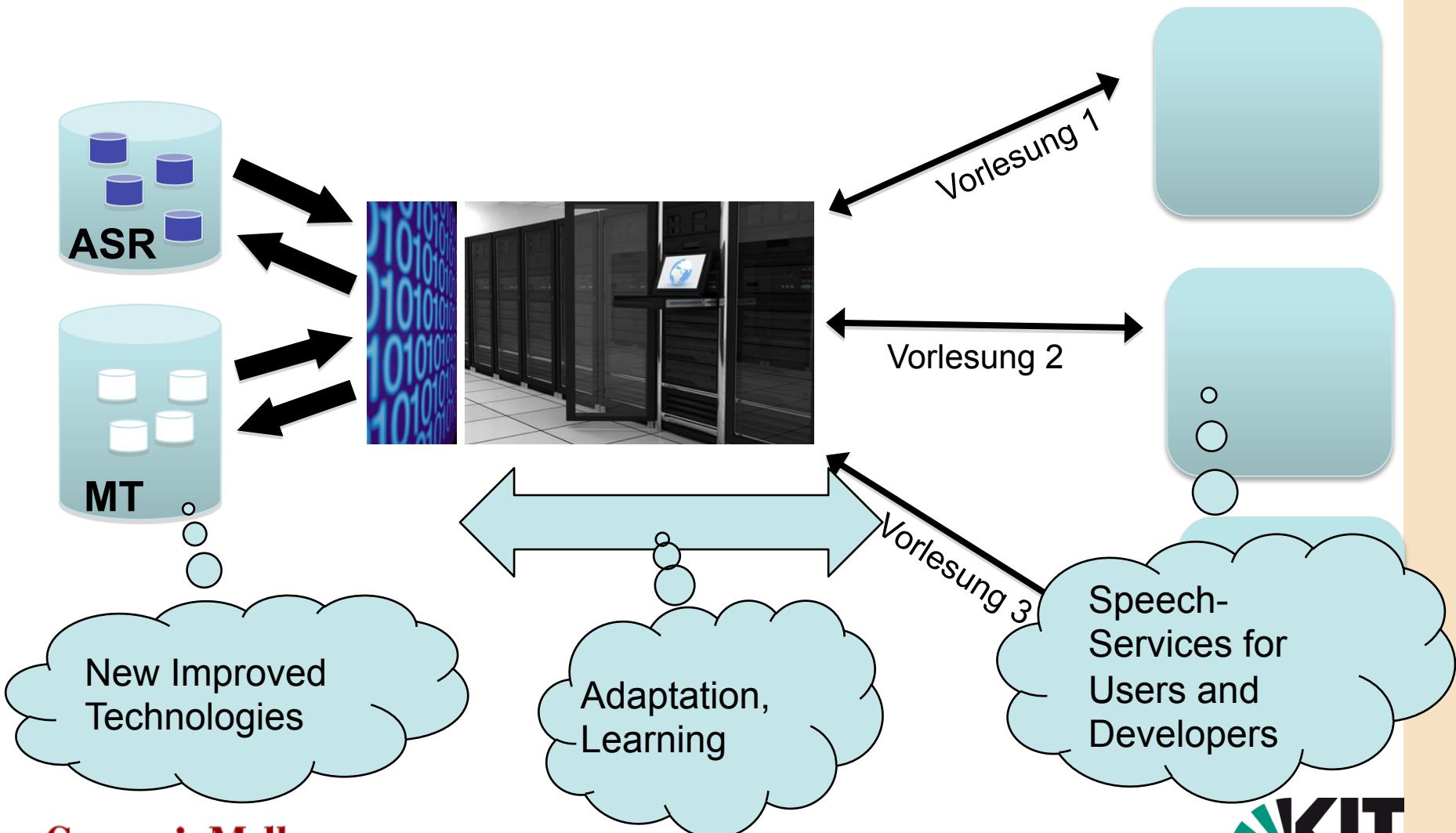
password: ShanghBJ

Service Infrastructure

Components

Services

Events



Mehr als nur Bahnhof verstehen - Weltweit erster Vorlesungsübersetzer

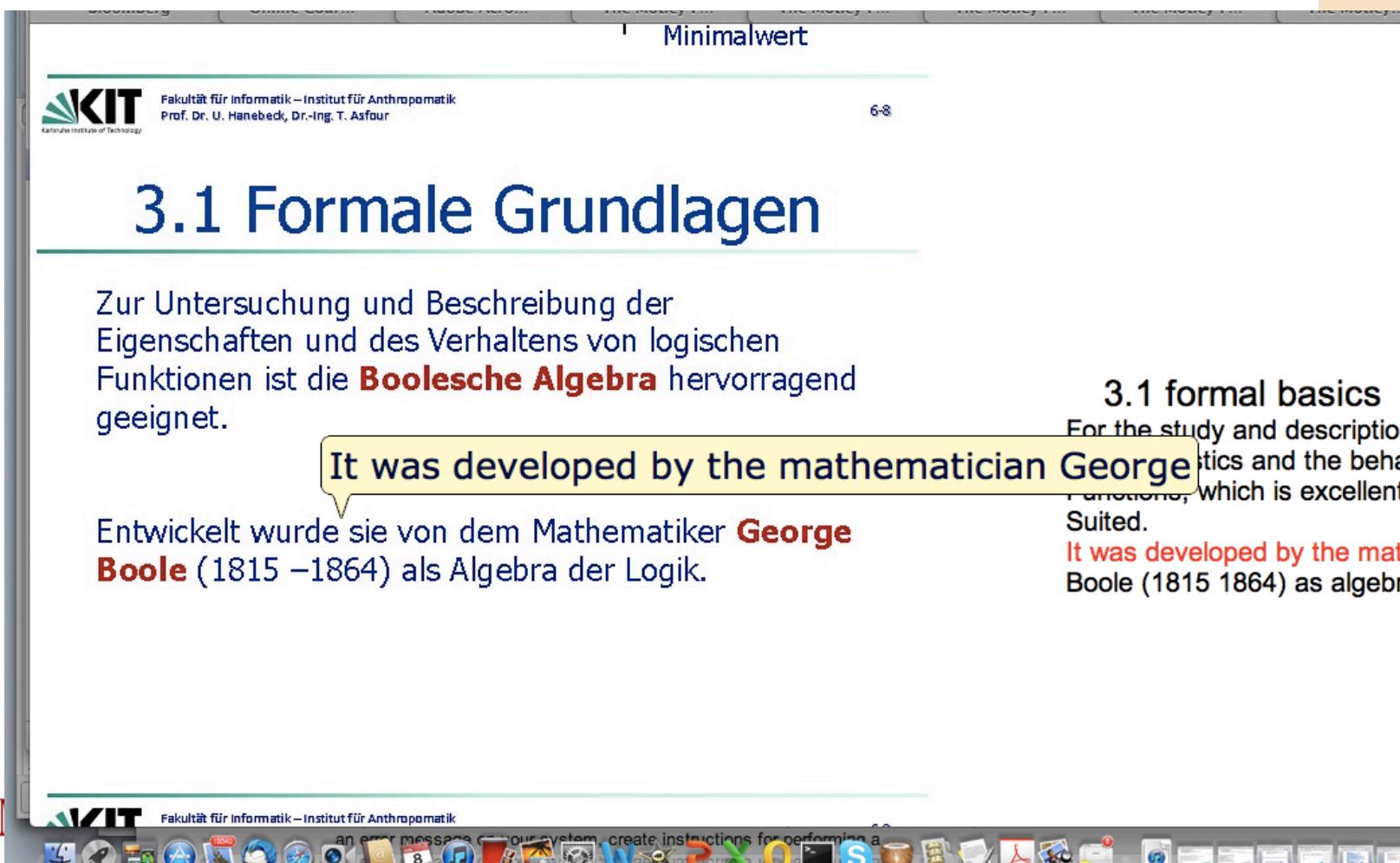
Karlsruhe. In einer deutschen Vorlesung sitzen und wegen der Sprachbarriere nur Ball verstehen - sowas könnte für ausländische Studenten bald Vergangenheit sein. Der w erste automatische Vorlesungsübersetzer ermöglicht Studierenden künftig, dem Vorti



Carnegie von Dozenten auf Englisch zu folgen - schriftlich übersetzt in Echtzeit. Am Montag Prof. Alex Waibel



- Translation of Powerpoint Slides
- Presentation by Subtitles



The screenshot shows a presentation slide with the following content:

Minimalwert

KIT
Karlsruhe Institute of Technology

Fakultät für Informatik – Institut für Anthropomatik
Prof. Dr. U. Hanebeck, Dr.-Ing. T. Asfour

6-8

3.1 Formale Grundlagen

Zur Untersuchung und Beschreibung der Eigenschaften und des Verhaltens von logischen Funktionen ist die **Boolesche Algebra** hervorragend geeignet.

It was developed by the mathematician George Boole (1815 –1864) als Algebra der Logik.

3.1 formal basics
For the study and description of the properties and the behavior of logic functions, which is excellent Suited.
It was developed by the mathematician George Boole (1815 –1864) as algebra of logic.

Carnegie

KIT

Fakultät für Informatik – Institut für Anthropomatik

An error message from your system, create instructions for performing a

- Transcripts useful to Search for Content
 - Slides, and Lectures in the Cloud
 - Search in Lectures and Foils by Way of Search Terms

LECTURE STORAGE

[Home](#) [Help](#) [Sign in](#)

List of Lectures

Fourier

Time	Lecture	Presenter
2012-05-24 14:24:29.544848	Kognitive Systeme - ro24052012	KIT-Dillmann
2012-04-23 14:49:12	KogSys 23.04.2012	KogSys DSP2 Waibel
2012-05-10 15:30:26	Rechner Organisation - RO100512	Asfour
2012-04-17 18:14:07	Rechnerorganisation_17_04_2012	dillmann and asfour are talking
2012-05-07 15:05:01	Kognitive Systeme	Mohr
2012-04-23 15:31:39	Kogsys 23.04.12	NULL
2012-05-09 13:02:38	Kognitive Systeme - KogSys 090512	Waibel
2012-05-23 13:07:13	Kognitive Systeme - kogsys230512	NULL
2012-05-23 13:07:51	Kognitive Systeme - Dozent: Prof. Waibel // Kognitive Systeme 23 05 2012	NULL
2012-05-31 11:32:36.774097	General - Jan MT 310512	NULL
2012-05-08 17:19:57	Kognitive Systeme - MT zweiter Teil	Waibel
2012-04-18 13:00:43	Kognitive Systeme 18.04.2012	NULL
2012-06-04 14:14:30.634523	Kognitive Systeme - ASR 3 und aktuelle Forschung	KIT-Waibel2

[previous](#) [next](#)

Simultaneous Translation of Lectures

- Continuous Monologue
 - Broadcast News, Speeches, Lectures
- Speaking-Style
 - Fast, spontaneous, fragmentary, and no punctuation!!
 - Noise, Caughing, Singing (!)
- Vocabulary
 - Much larger, Special Vocabularies
- Speed, Realtime
- Service-Infrastructure
 - Many parallel lectures;
 - Automatic, robust assignment of compute power

- MT in **German** Lectures is particularly hard. Why?
- Peculiarities of German:
 - Wordorder:

Ich möchte mich zu der Konferenz über Maschinelle Übersetzung *anmelden*
→ I want to register to the conference on Machine Translation
 - Compounds:

Worterkennungsfehlerrate
→ Word Recognition Error Rate
 - Inflections and Agreement:

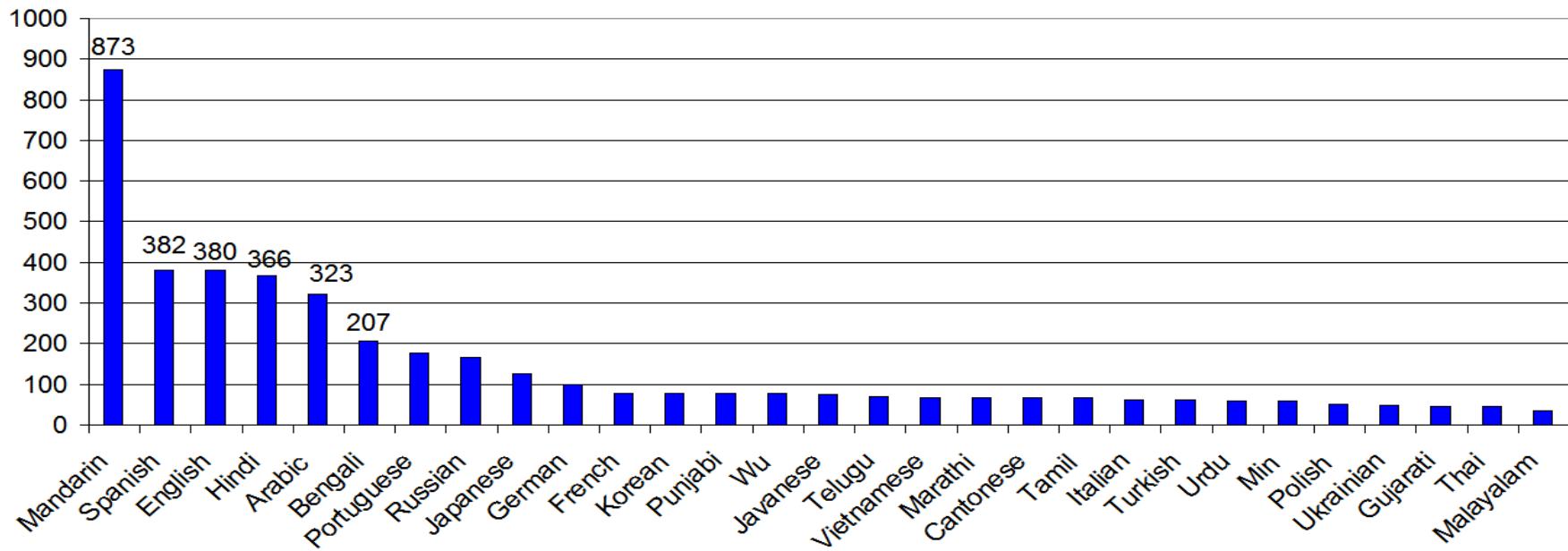
Zu der nächsten wichtigen interessanten Vorlesung

- Technical Terms
normally not in ‘normal’ vocabularies
 - Cepstral-Koeffizienten
 - Wälzlagerungen → Roller Bearings
 - Unterraum → Subspace
- Technical Terms
with special Meanings
 - Klausur → Final Exam (not Retreat)
 - Vorzeichen → Sign (not Omen)
- Formulas:
 - Eff von Ix → $f(x)$

- Foreign Words in German Language
 - Computer Science, English Expressions
 - Political Speeches, Latin Proverbs
- Accent
 - “Würfelkalkül” (Asfour)
- Foreign Words in German Language
 - “Cloud”, “iPhone”, “iPad”, “Laser”
- Inflection und Declination of these Words
 - Web-ge-casted, down-ge-loaded
- Formation of Compounds:
 - Cloudbasierter Webcastzugriff

- Languages:
 - Only a Few Languages are Currently Addressed (<10)
 - Development of Technology Takes Long & Is Expensive
 - Cost more than 1M \$ and DevTime more than 1 Year per Language
 - 6, 000 languages, 36 M potential language pairs, Plus Dialects
 - Technology is Always a Step (or Two!) Behind Deployment

Languages by Million Native Speakers



Conclusion

Communication between the people of the world

- Bridging the Linguistic Divide
- Methods: Machine Learning from Data
- User Interaction
- Adaptation, Learning, Forgetting
- New Possibilities for Broader Class of Application
- Better Language Portability
- More Robust Performance
- Integration into Services

