



The first International Workshop on Spoken Languages Technologies for Under-resourced languages (SLTU - 2008)

Invited talk 1

Rapid Language Adaptation Tools and Technologies for Multilingual Speech Processing Systems

Tanja Schultz

University of Karlsruhe and Carnegie Mellon University

The performance of speech and language processing technologies has improved dramatically over the past decade, with an increasing number of systems being deployed in a large variety of applications, such as spoken dialog systems, speech summarization and information retrieval systems, and speech translation systems. Most efforts to date were focused on a very small number of languages with large number of speakers, economic potential, and information technology needs of the population. However, speech technology has a lot to contribute even to those languages that do not fall into this category. Languages with a small number of speakers and few linguistic resources may suddenly become of interest for humanitarian and military reasons. Furthermore, a large number of languages are in danger of becoming extinct, and ongoing projects for preserving them could benefit from speech technology.

With more than 6900 languages in the world and the need to support multiple input and output languages, the most important challenge today is to port speech processing systems to new languages rapidly and at reasonable costs. Major bottlenecks are the lack of data and language conventions, and the gap between technology and language expertise. The lack of data results from the fact that today's speech technologies heavily rely on statistically based modeling schemes, such as Hidden Markov Models and n-gram language modeling. Although statistical modeling algorithms are mostly language independent and proved to work well for a variety of languages, the parameter estimation requires vast amounts of training data. Large-scale data resources are currently available for less than 50 languages and the costs for these collections are prohibitive to all but the most widely spoken and economically viable languages. In addition, a surprisingly large number of languages or dialects lack a standardized writing system which hinders web harvesting of large text corpora or the construction of dictionaries and lexicons. Last but not least, despite the well-defined process of system building it is very cost- and time consuming to handle language-specific peculiarities, and it requires substantial language expertise. Unfortunately, it is extremely difficult to find system developers who simultaneously have the necessary technical background and significant insight into the language in question. Consequently, one of the central issues in developing systems in many input and output languages is the challenge of bridging the gap between language and technology expertise.

In my talk I will introduce state-of-the-art techniques for rapid language adaptation and present existing solutions to overcome the ever-existing problem of data sparseness and the gap between language and technology expertise. I will describe the building process for speech recognition and speech synthesis components for new unsupported languages and introduce tools to do this rapidly and at low costs. The talk describes the SPICE Toolkit (Speech Processing - Interactive Creation and Evaluation), a web based toolkit for rapid language adaptation to new languages. The methods and tools implemented in SPICE enables user to develop speech processing components, to collect appropriate data for building these models, and to evaluate the results allowing for iterative improvements. Building on existing projects like GlobalPhone and FestVox, knowledge and data are shared between recognition and synthesis; this includes phone sets, pronunciation dictionaries, acoustic models, and text resources. SPICE is an online service (<http://cmuspice.org>). By archiving the data gathered on-the-fly from many cooperative users we hope to significantly increase the repository of languages and resources and make the data and components for new languages available at large to the community. By keeping the users in the developmental loop, SPICE tools can learn from their expertise to constantly adapt and improve. This will hopefully revolutionize the system development process for new languages.

The first International Workshop on Spoken Languages Technologies for Under-resourced languages (SLTU - 2008)